

Research Article

Probabilistic Ensemble Forecast of Summertime Temperatures in Pakistan

Muhammad Hanif

Pakistan Meteorological Department, National Weather Forecasting Centre, P.O. Box 1214, Islamabad 44000, Pakistan

Correspondence should be addressed to Muhammad Hanif; hanifwxc@hotmail.com

Received 9 November 2013; Accepted 28 February 2014; Published 30 March 2014

Academic Editor: Harry D. Kambezidis

Copyright © 2014 Muhammad Hanif. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Snowmelt flooding triggered by intense heat is a major temperature related weather hazard in northern Pakistan, and the frequency of such extreme flood events has increased during the recent years. In this study, the probabilistic temperature forecasts at seasonal and subseasonal time scales based on hindcasts simulations from three state-of-the-art models within the DEMETER project are assessed by the relative operating characteristic (ROC) verification method. Results based on direct model outputs reveal significant skill for hot summers in February 3–5 (ROC area = 0.707 with lower 95% confidence limit of 0.538) and February 4–5 (ROC area = 0.771 with lower 95% confidence limit of 0.623) forecasts when validated against observations. Results for ERA-40 reanalysis also show skill for hot summers. Skilful probabilistic ensemble forecasts of summertime temperatures may be valuable in providing the foreknowledge of snowmelt flooding and water management in Pakistan.

1. Introduction

Skilful seasonal forecasts of temperature have obvious potential benefits, particularly in developing countries, where agriculture and water resources are the main sectors directly affected by temperature changes and climate variability. More efficient use of available resources is possible if a reliable forecast of temperatures is available. The ensemble based probabilistic forecasting approach has become an integral part of seasonal predictions, and some meteorological ensemble prediction systems (EPS) are operational at a global scale [1, 2]. In a European funded project, DEMETER, state-of-the-art coupled models have the capability to correctly simulate the rainfall and temperatures as mentioned in several studies. The multimodel ensemble system used in the DEMETER has shown the skill for seasonal forecasting of maximum temperature for different seasons in Spain [3]. The hindcasts produced within the DEMETER project have shown good skills for the winter season over northern Italy [4]. Weigel et al. [5] also showed that a single recalibrated model of the DEMETER project can also perform nearly equal to multimodel ensemble predictions. Krishnamurti et al. [6] compared the skills from the DEMETER dataset and a suite

of Florida State University (FSU) models and suggested that a superensemble of multiple coupled models yields considerable skill for different variables, including precipitation and temperatures, of the south Asian monsoon.

DEMETER provides substantial information concerning the quality of predictions from numerical models that there is as yet no equivalent project for detailing the efficacy of empirical models [7]. The DEMETER database has not been widely used to examine the skills for seasonal forecasts for the South Asian region where precipitation and temperatures have direct and serious societal impacts. The snowmelt flooding is rare in the world and the prediction is therefore not common. Northern Pakistan is one of those few regions of the world which is affected by snowmelt flooding events. Heavy winter snow cover, rapid snow melt, and heavy rains during the melt period are the major factors that contribute to snowmelt flooding. The major rivers of Pakistan have their origin in the higher altitudes and derive their flows mainly from snowmelt and monsoon rains. The months of July and August are the peak rainy seasons of monsoon in Pakistan and concentrated rainfall in the catchments during these months is sometimes amplified by snowmelt flows. Minor snowmelt flooding occurs every year in northern Pakistan

that affects the localized areas but sometimes extreme heat triggers large scale snowmelt leading to high flooding in the rivers. Snowmelt runoff leading to flash floods, in northern Pakistan, is regulated by summer temperatures [8]. Extremely high temperatures over northern Pakistan cause snowmelt floods which often take the shape of a disaster that disrupts irrigation and people's lives in mountainous areas. The intense heat caused several extreme flood events in the past leading to dangerously high river levels in Pakistan. After a disastrous snowmelt flood event in June 2005, the demand for flood alerts based on an accurate prediction of summertime temperatures has become an issue of concern. No attempt in the past, unfortunately, has been undertaken to predict summer temperatures, particularly in northern Pakistan. This knowledge gap is the main motivation of this study, and the DEMETER multimodel ensemble prediction system has been used for the prediction of summer temperatures at seasonal and subseasonal time scales in northern Pakistan. Subseasonal predictability is equally important, particularly in the context of weather and climate risk management [9], but limited literature has been published about it. ECMWF (European Centre for Medium Range Weather Forecast) and JMA (Japan Meteorological Agency) are major international weather centres that operationally issue monthly forecasts on the basis of a global circulation model [10, 11]. Frías et al. [12] described the predictability of monthly average maximum temperatures by multimodel ensemble for seasonal to interannual prediction system (DEMETER) over the Iberian Peninsula. After presenting a probabilistic verification of monthly temperature forecasts using a coupled atmosphere-ocean global circulation model of ECMWF, Weigel et al. [9] have suggested that monthly forecasting bridges the gap between medium-range weather forecasting and seasonal predictions.

The paper is structured as follows. Section 2 describes the datasets and methods used to assess the probabilistic ensemble forecasts. In Section 3, the key results of probabilistic ensemble forecasts of temperatures are described. Finally, the main results and conclusions derived in this study are summarized in Section 4.

2. Data and Method

This paper aims to investigate temperature predictability during the summer season in northern Pakistan. Several cities are located in the mountainous region of northern Pakistan, and the narrow and steep valleys of the region cause a fast reaction to snowmelt flood waves. The DEMETER, ERA-40 reanalysis, and observational datasets are used in this study for investigating the seasonal reforecasts, assessed by ROC verification method.

The DEMETER system comprises seven state-of-the-art global coupled ocean-atmosphere models, used to produce an extensive hindcasts dataset starting from February, May, August, and November, for different periods ranging from 1958 to 2001 [13]. Three of the seven coupled models (CNRM—from Centre National de Recherches Météorologiques, ECMWF—from European Centre for

Medium-Range Weather Forecasts, and UKMO—from Met Office, UK) within DEMETER produced hindcasts for a long period of 43 years (1959–2001), and these three models have been used in this study to assess the seasonal temperatures over northern Pakistan. In the DEMETER project, atmospheric and land-surface initial conditions are taken from the ERA-40 project [14]. The DEMETER hindcasts starting in February have been examined in predicting the summertime temperatures at seasonal and subseasonal time scales. The observational data of mean monthly and seasonal temperatures averaged over northern Pakistan, used for forecasts validation, was provided by the climate data processing centre (CDPC) of Pakistan Meteorological Department (PMD).

The study consists of probabilistic forecasts of temperatures considering three categories (hot/above-the-median/cold), and the performance of forecasts has been assessed using the relative operating characteristic (ROC) verification method [15]. ROC curves are used to illustrate the forecast quality, and the first way to quantify the ROC is to calculate the area beneath the ROC curve [16]. Area under the curve is nonparametrical, is used for measuring the test accuracy, and is also used to compare the shapes of their corresponding curves.

A major limitation of the multimodel ensemble approach is the influence of systematic errors in the coupled global circulation models (GCMs) outputs, and the most sophisticated coupled GCMs may have inherent deficiencies which are often shared between models [17]. The model errors (biases) can be removed by calibration against observations, and the simplest possible method of calibration is bias-correction. The resulting ensemble is generally more accurate if each ensemble member is bias-corrected [18]. A simple postprocessing scheme for the bias-corrected ensemble system can be used to calculate the reasonably accurate predictions [19, 20]. The ensemble forecast products may enhance significantly after climate mean bias-correction [21]. The main focus of the study is on prediction skills by the direct model outputs (raw). However, the forecast skills have also been examined after bias-correction, and the ERA-40 reanalysis was used for this bias-correction. Bias-correction was applied to each ensemble member of each global coupled model (GCM) separately. So two sets of input ensemble data (uncorrected and bias-corrected) have been used to examine and compare the temperatures predictability in Pakistan from DEMETER.

3. Results

The computed ROC areas (uncorrected and bias-corrected) for three temperature categories (upper tercile (UT), above-the-median (AM), and lower tercile (LT)) in February (3–5) forecasts, validated against observations and ERA-40, are given in Table 1. For the upper tercile category (hot summer), results showed significant skill against observations (ROC area = 0.707 with lower 95% confidence limit of 0.538) and marginal skill against ERA-40 reanalysis (ROC area = 0.656 with lower 95% confidence limit of 0.484). Results also revealed marginal skills for AM and LT categories. After bias-corrections, the skills for all categories have slightly

TABLE 1: ROC areas for February (3–5) forecasts by DEMETER-3 validated against observations and ERA-40 data over northern Pakistan. The range of values in brackets represents 95% CI calculated with 1000 bootstrap samples. ROC areas >0.5 are skillful and <0.5 are unskillful. Skillful forecasts are highlighted in bold values.

February (3–5) forecasts	ROC areas against observations		ROC areas against ERA-40	
	Uncorrected	Bias-corrected	Uncorrected	Bias-corrected
Upper tercile (UT) (95% CI)	0.707 (0.538–0.876)	0.727 (0.563–0.891)	0.656 (0.484–0.828)	0.690 (0.527–0.852)
Above-the-median (AM) (95% CI)	0.623 (0.455–0.791)	0.687 (0.527–0.848)	0.623 (0.455–0.791)	0.667 (0.505–0.828)
Lower tercile (LT) (95% CI)	0.637 (0.470–0.803)	0.651 (0.486–0.817)	0.675 (0.512–0.838)	0.708 (0.548–0.869)

TABLE 2: ROC areas for February (4–5) forecasts by DEMETER-3 validated against observations and ERA-40 data over northern Pakistan.

February (4–5) forecasts	ROC areas against observations		ROC areas against ERA-40	
	Uncorrected	Bias-corrected	Uncorrected	Bias-corrected
Upper tercile (UT) (95% CI)	0.771 (0.623–0.919)	0.780 (0.642–0.918)	0.675 (0.510–0.840)	0.683 (0.523–0.844)
Above-the-median (AM) (95% CI)	0.675 (0.514–0.837)	0.687 (0.529–0.845)	0.639 (0.471–0.806)	0.641 (0.474–0.807)
Lower tercile (LT) (95% CI)	0.569 (0.391–0.747)	0.601 (0.428–0.774)	0.536 (0.365–0.706)	0.559 (0.390–0.728)

increased. ROC curves for hot summer are displayed in Figure 1, which also illustrate significant skills for these events. The temperature skills at the subseasonal time scale were also examined, and the results, given in Table 2, revealed significant skills for hot summers (UT) in February (4–5) forecast when validated against observations (ROC area = 0.771) and ERA-40 (ROC area = 0.675). The bias-corrected results showed a positive impact on temperature forecast skills. ROC curves for subseasonal time scales are given in Figure 2, and the shapes of curves also illustrate high skill for the upper tercile category (hot summer).

The individual forecast probabilities of the upper tercile (UT) from February (4–5) forecasts by DEMETER-3 multimodel ensembles, for 1959–2001, have been examined in Figure 3. The plotted forecast probabilities of the UT category for individual years reveal that DEMETER-3 has predicted the year 1967 as upper tercile with the highest forecast probability, which is not an upper tercile year according to the observations of PMD. However, the probability plots for 43 years (1959–2001) showed that the DEMETER-3 has predicted several UT years correctly (i.e., 10 out of 14 years correctly predicted) when high discharge was observed at Attock (where snowmelt water of both River Indus and River Kabul meets). This probability plot confirmed that February (4–5) forecasts by DEMETER-3 have high skill in predicting the UT events. The individual forecast probabilities of upper tercile (UT) from February (3–5) forecasts also revealed the same level of skill (not shown).

4. Discussion and Conclusions

The DEMETER forecasts starting in February for summertime temperatures in northern Pakistan have been assessed by the ROC verification method. Significant skill is found for

hot summers in February (3–5) and February (4–5) forecasts when validated against observational data of PMD and ERA-40 reanalysis data. Skillful probabilistic ensemble forecast of summer temperatures in northern Pakistan is extremely valuable in flood management, and the use of probabilistic forecast by DEMETER may be a better way for risk-based decision-making in a flood warning. Sensitivity of flood events to summertime temperatures should be examined carefully by further research to develop an appropriate relationship between summertime temperature input and runoff response for flood risk. Such a relationship will be helpful for decision-makers for prior assessment of flood intensity. Due to limited water storage capacity in Pakistan, a huge amount of snowmelt water is wasted during dry summer months (April–June). Pakistan is already one of the most water-stressed countries in the world due to high population growth [22]. The prior knowledge of snowmelt water, based on skillful probabilistic ensemble forecasts of summer temperatures, may be helpful for proper planning and improving the water management system in Pakistan.

The predictability of hot and mild summers has been discussed in several studies but the prediction of snowmelt triggered by extreme temperatures is unfortunately negligible because limited areas in the world are affected by snowmelt flooding. The snowmelt is a major source of fresh water in the world which is mostly wasted in developing countries like Pakistan. The models used in the DEMETER project or ENSEMBLES have the potential to predict the extreme high temperatures of a smaller area like northern Pakistan at seasonal and subseasonal scales to alert the policy makers and the stakeholders about potential snowmelt flooding. Such information can also be used by the hydrologists in snowmelt modelling and it may provide precious information for the forecasters to remain alert regarding snowmelt flooding.

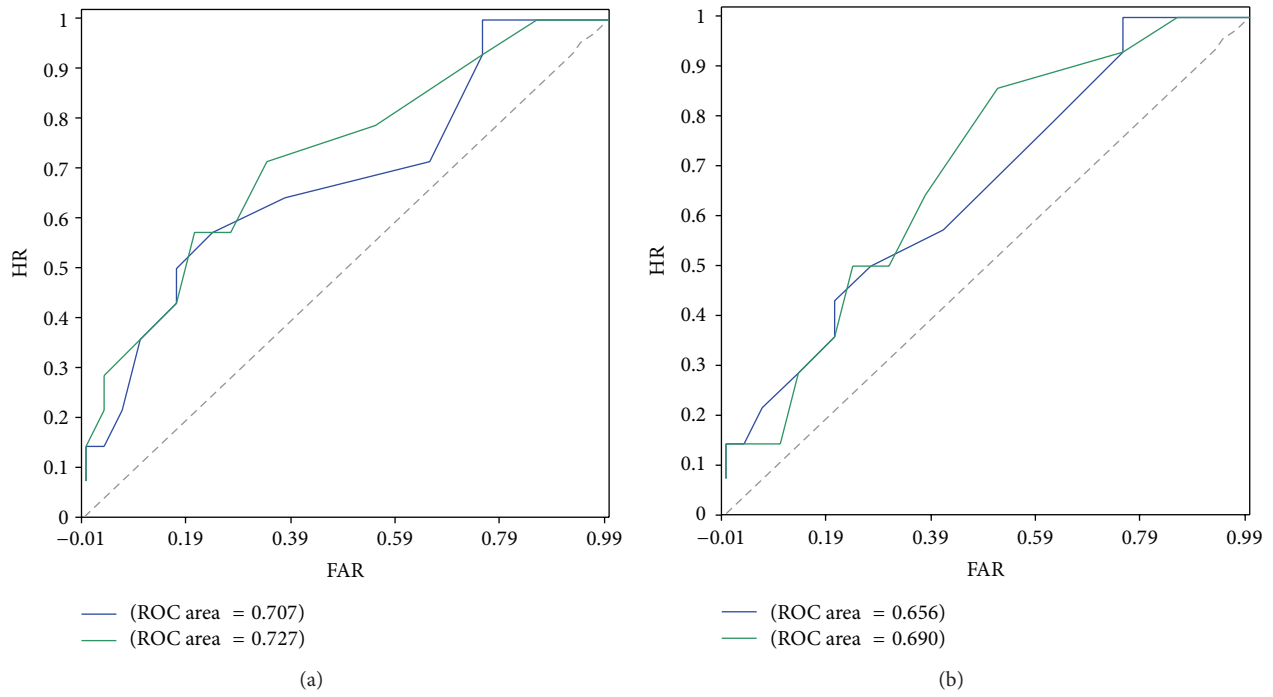


FIGURE 1: ROC curves of February (3–5) forecasts for UT categories using the DEMETER-3 ensemble over northern Pakistan for the period of 1959–2001 and validated against (a) observations and (b) ERA-40. Blue and green lines represent uncorrected and bias-corrected skills, respectively.

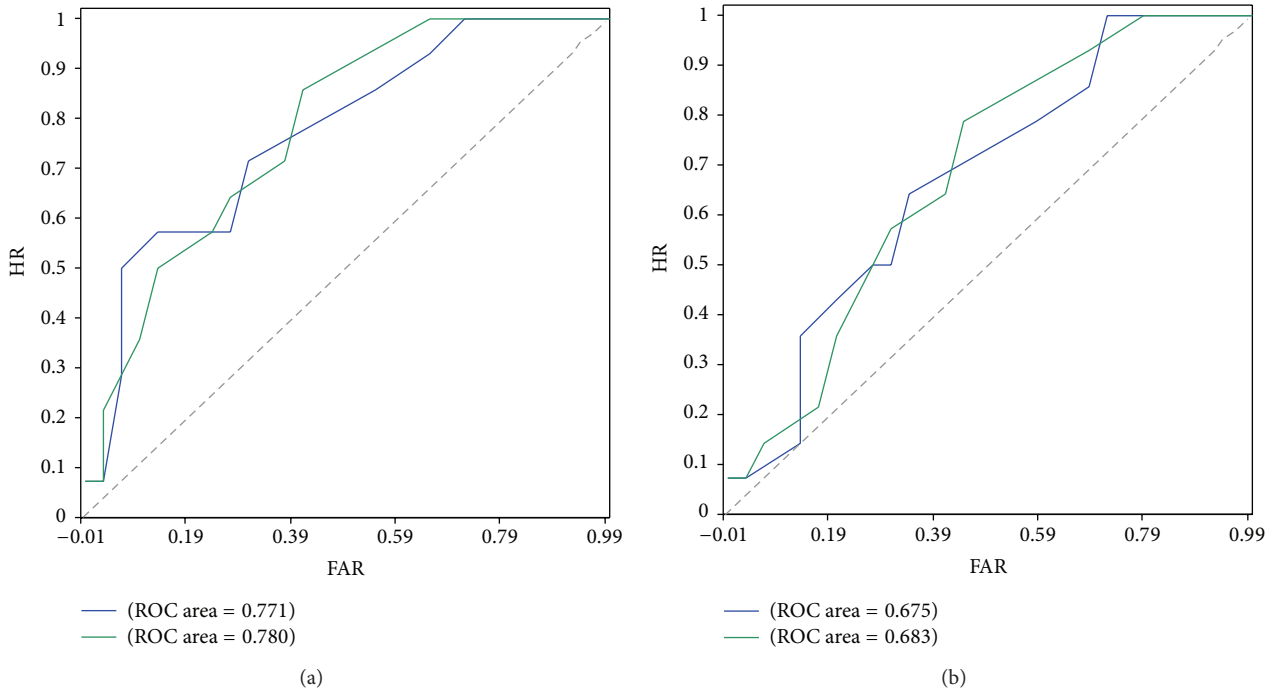


FIGURE 2: ROC curves of February (4–5) forecasts for UT categories using the DEMETER-3 ensemble over northern Pakistan for the period of 1959–2001 and validated against (a) observations and (b) ERA-40. Blue and green lines represent uncorrected and bias-corrected skills, respectively.

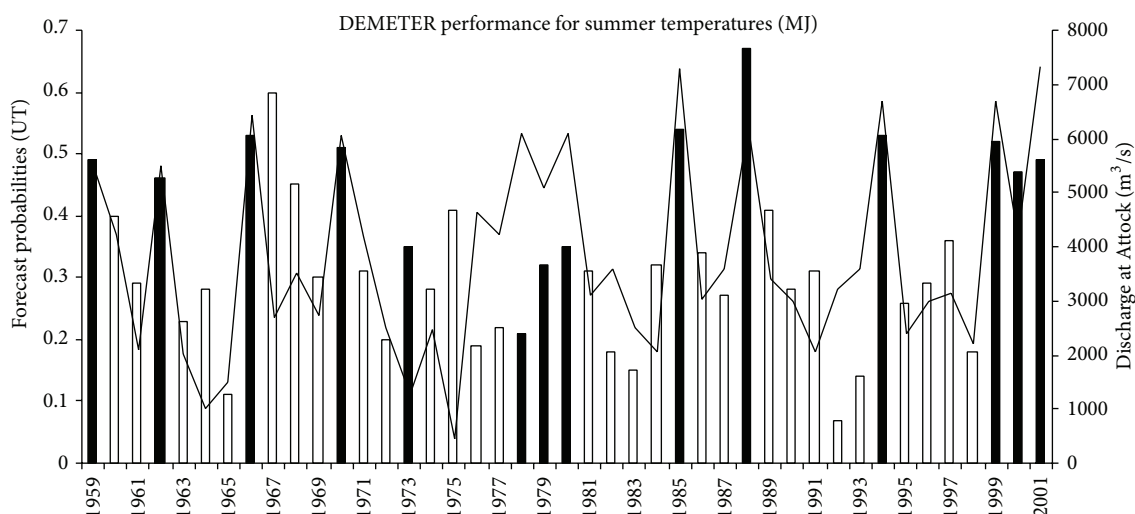


FIGURE 3: DEMETER forecast probabilities of upper tercile (UT) temperature obtained from February (4-5) forecast in northern Pakistan for the period of 1959–2001. Upper tercile years according to the observational data of PMD are shown with shaded bars and other years are shown with hollow bars. The line-plot indicates the average discharge data of May and June at Attock (River Indus).

The temperature forecasts skill, based on direct model outputs, may further be improved using the appropriate statistical downscaling approach, feasible in northern Pakistan, a region of complex topography. The potential probabilistic temperature forecasts based on a multimodel ensemble prediction system as that used in the DEMETER (the precursor of the operational EURO-SIP model) may have enormous socioeconomic benefits for Pakistan, Afghanistan, and other areas of central south west Asia which are prone to heavy snow in winter and snowmelt flooding during the spring and summer for better water management. The same three models of the DEMETER have subsequently been used in the EUROSIP project to produce operational multimodel seasonal forecasts.

Conflict of Interests

The author declares that there is no conflict of interests regarding the publication of this paper.

References

- [1] R. Buizza, P. L. Houtekamer, Z. Toth, G. Pellerin, M. Wei, and Y. Zhu, "A comparison of the ECMWF, MSC, and NCEP global ensemble prediction systems," *Monthly Weather Review*, vol. 133, no. 5, pp. 1076–1097, 2005.
- [2] T. Gneiting and A. E. Raftery, "Weather forecasting with ensemble methods," *Science*, vol. 310, no. 5746, pp. 248–249, 2005.
- [3] M. D. Frías, S. Herrera, A. S. Cofiño, and J. M. Gutiérrez, "Assessing the skill of precipitation and temperature seasonal forecasts in Spain: windows of opportunity related to ENSO events," *Journal of Climate*, vol. 23, no. 2, pp. 209–220, 2010.
- [4] V. Pavan, S. Marchesi, A. Morgillo, C. Cacciamani, and F. J. Doblas-Reyes, "Downscaling of DEMETER winter seasonal hindcasts over Northern Italy," *Tellus A: Dynamic Meteorology and Oceanography*, vol. 57, no. 3, pp. 424–434, 2005.
- [5] A. P. Weigel, M. A. Liniger, and C. Appenzeller, "Seasonal ensemble forecasts: are recalibrated single models better than multimodels?" *Monthly Weather Review*, vol. 137, no. 4, pp. 1460–1479, 2009.
- [6] T. N. Krishnamurti, A. K. Mitra, T. S. V. Vijaya Kumar, W. T. Yun, and W. K. Dewar, "Seasonal climate forecasts of the South Asian monsoon using multiple coupled models," *Tellus A: Dynamic Meteorology and Oceanography*, vol. 58, no. 4, pp. 487–507, 2006.
- [7] M. Harrison, "The development of seasonal and inter-annual climate forecasting," *Climatic Change*, vol. 70, no. 1-2, pp. 201–220, 2005.
- [8] D. R. Archer and H. J. Fowler, "Using meteorological data to forecast seasonal runoff on the River Jhelum, Pakistan," *Journal of Hydrology*, vol. 361, no. 1-2, pp. 10–23, 2008.
- [9] A. P. Weigel, D. Baggenstos, M. A. Liniger, F. Vitart, and C. Appenzeller, "Probabilistic verification of monthly temperature forecasts," *Monthly Weather Review*, vol. 136, no. 12, pp. 5162–5182, 2008.
- [10] F. Vitart, "Monthly forecasting at ECMWF," *Monthly Weather Review*, vol. 132, no. 12, pp. 2761–2779, 2004.
- [11] T. Tokioka, "Chapter 12 Climate services at the Japan meteorological agency using a general circulation model: dynamical one-month prediction," *International Geophysics*, vol. 70, no. C, pp. 355–371, 2001.
- [12] M. D. Frías, J. Fernández, J. Sáenz, and C. Rodríguez-Puebla, "Operational predictability of monthly average maximum temperature over the Iberian Peninsula using DEMETER simulations and downscaling," *Tellus A: Dynamic Meteorology and Oceanography*, vol. 57, no. 3, pp. 448–463, 2005.
- [13] T. N. Palmer, A. Alessandri, U. Andersen et al., "Development of a European multimodel ensemble system for seasonal-to-interannual prediction (DEMETER)," *Bulletin of the American Meteorological Society*, vol. 85, no. 6, pp. 853–872, 2004.
- [14] S. Uppala, P. W. Kallberg, A. J. Simmons et al. et al., "The ERA-40 re-analysis," *Quarterly Journal of the Royal Meteorological Society*, vol. 131, no. 612, pp. 2961–3012, 2005.

- [15] I. T. Jolliffe and D. B. Stephenson, *Forecast Verification: A Practitioner's Guide in Atmospheric Science*, John Wiley & Sons, New York, NY, USA, 2003.
- [16] D. M. Green and J. A. Swets, *Signal Detection Theory and Psychophysics*, John Wiley & Sons, New York, NY, USA, 1966.
- [17] C. Brankovic and T. N. Palmer, "Seasonal skill and predictability of ECMWF PROVOST ensembles," *Quarterly Journal of the Royal Meteorological Society*, vol. 126, no. 567, pp. 2035–2067, 2000.
- [18] A. E. A. Hamm, "Validation of the NCEP SREF," in *Proceedings of the 20th AMS Conference on Weather Analysis and Forecasting and 16th AMS Conference on Numerical Weather Prediction*, American Meteorological Society, Seattle, Wash, USA, 2003.
- [19] N. Yussouf and D. J. Stensrud, "Prediction of near-surface variables at independent locations from a bias-corrected ensemble forecasting system," *Monthly Weather Review*, vol. 134, no. 11, pp. 3415–3424, 2006.
- [20] W. Kleiber, A. E. Raftery, J. Baars, T. Gneiting, C. F. Mass, and E. Gritmit, "Locally calibrated probabilistic temperature forecasting using geostatistical model averaging and local bayesian model averaging," *Monthly Weather Review*, vol. 139, no. 8, pp. 2630–2649, 2011.
- [21] B. Cui, Z. Toth, Y. Zhu, and D. Hou, "Bias correction for global ensemble forecast," *Weather and Forecasting*, vol. 27, no. 2, pp. 396–410, 2012.
- [22] World Bank, "Pakistan country water resources assistance strategy, water economy: running dry," Tech. Rep. 34081-PK, 2005, <http://siteresources.worldbank.org/PAKISTANEXTN/Resources/PWCAS-Title&Intro.pdf>.

