Supportive empirical modelling for the forecast of monsoon precipitation in Nepal

ARTICLE *in* INTERNATIONAL JOURNAL OF CLIMATOLOGY · NOVEMBER 2013

Impact Factor: 3.16 · DOI: 10.1002/joc.3649

4 AUTHORS, INCLUDING:

Robert Gillies
Utah State University

Shih-Yu Simon Wang
Utah State University / Utah Climate Center

Yan Sun
Utah State University

INTERNATIONAL JOURNAL OF CLIMATOLOGY
Int. J. Climatol. (2013)
Published online in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/joc.3649

Supportive empirical modelling for the forecast of monsoon precipitation in Nepal

Robert R. Gillies,[a,b]* Shih-Yu Wang,[a,b] Yan Sun[c] and Oi-Yu Chung[a,b]
[a] *Utah Climate Center, Utah State University, Logan, UT, USA*
[b] *Department of Plants, Soils, and Climate, Utah State University, Logan, UT, USA*
[c] *Department of Mathematics and Statistics, Utah State University, Logan, UT, USA*

ABSTRACT: Seasonal prediction of the monsoon precipitation in Nepal has been a challenge. That is partly because Nepal's monsoon precipitation exhibits a distinct and strong quasi-decadal oscillation while not correlated with the El Nino-Southern Oscillation. The existing global and regional climate models are insufficient in deriving reliable precipitation prediction. This paper examines the prediction of Nepal's July to August (JA) mean precipitation using five different methods: three time series models in comparison with a persistence forecast (PF) and a climatology forecast (CF). The first model (P-AR) uses past precipitation data to forecast the future, based upon the recently uncovered quasi-periodic feature of the JA mean precipitation. The other two models (ARX-SST and ARX-GQ) add covariate sea surface temperature (SST) and global water vapour flux circulation (GQ) respectively. Based upon the evaluation of 1-year-ahead forecast, the three time series models performed better than PF and CF. Of those, the P-AR model has the least mean absolute error (MAE) of <1 mm day^{-1}. Based upon the 2-year-ahead forecast results, the P-AR model performs slightly better than ARX-SST and ARX-GQ models. The forecast ability of the time series models appears better than that of operational numerical models such as the NCEP Climate Forecast System (CFS) and so, can be used as an effective alternative in predicting monsoon precipitation for Nepal.

KEY WORDS Nepal monsoon; precipitation extreme; statistical model; prediction

Received 5 April 2012; Revised 1 December 2012; Accepted 9 December 2012

1. Introduction

The challenge that exists in the seasonal prediction of the monsoon precipitation in Nepal is threefold: (1) the precipitation there is uncorrelated with the 'all-India' monsoon precipitation (Shrestha *et al.*, 2000) and this makes the prediction techniques developed for the Indian monsoon inapplicable for Nepal, (2) a marked quasi-decadal oscillation (QDO) characterizes the monsoon precipitation in Nepal but not the 'all-India' precipitation (Wang and Gillies, 2012, herein WG12) and (3) the concentrated and highly convective monsoon precipitation along the Nepal Himalayas is difficult to simulate with climate models (Shrestha and Karmacharya, 2006). For instance, Pattanaik and Kumar (2010) showed that the NCEP Climate Forecast System (CFS; Saha *et al.*, 2006) has limited skill in predicting the monsoon precipitation in Nepal even with a short lead time ($\leq$2 months). Pattanaik and Kumar (2010) further pointed out two notable precipitation biases of the CFS in Nepal in that: (1) it severely over-predicted monsoon precipitation as a result of predominantly strong low-level convergence along the Nepal-Gangetic Plains, (2) there exists weak to negative correlations with verifying precipitation data. Although the former bias may be reduced by applying an applicable downscaling technique (Pattanaik *et al.*, 2010), the latter bias in variability remains unresolved.

Empirical prediction schemes of the Indian monsoon have been developed for over a century (Shukla and Mooley, 1987; Navone and Ceccatto, 1994; DelSole and Shukla, 2002; among others) and the majority of the prediction schemes are based on the El Niño-Southern Oscillation (ENSO; Yang *et al.*, 2008; Achuthavarier *et al.*, 2011). However, the monsoon precipitation in Nepal is uncorrelated with the Indian monsoon (Shrestha *et al.*, 2000); this is because the Indian monsoon variability is predominantly driven by ENSO whereas the correlation between the Nepal monsoon and ENSO is weak (WG12). In fact, as is illustrated in Figure 1, the cross correlations between the July to August (JA) mean precipitation in Nepal and the overlapping 2-month averages of the Nino-3, Nino-3.4 and Nino-4 indices from JA to the November to December (ND) of the previous year are weak. [The indices are sea surface temperature anomalies averaged within the Nino-3.4 (170°W–120°W), Nino-4 (160°E–150°W), and Nino-3 (210°W–270°W) domains between 5°S–5°N, based on the ERSST data (see Section 2); precipitation is averaged within the geographical boundary of Nepal.] Moreover,

* Correspondence to: Robert R. Gillies, Utah Climate Center, Utah State University, Logan, UT, USA. E-mail: Robert.Gillies@usu.edu

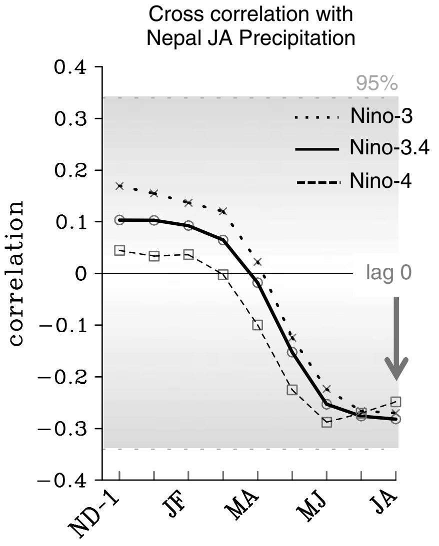

Figure 1. Cross correlations between July to August (JA) mean precipitation in Nepal and the overlapping 2-month averages of Nino-3.4, Nino-3 and Nino-4 indices from preceding November to December through JA, 1951–2007. The shaded region indicates the range of the 95% significance level.

when it comes to predicting the monsoon with the CFS, Yang *et al.* (2008) noted that the CFS tends to overemphasize the relationship between ENSO and the Indian monsoon and, this further impedes the CFS's skill in predicting the monsoon precipitation in Nepal. Given such circumstances, a statistical approach that combines precipitation and other climate variables may yield reasonable skill in predicting Nepal's monsoon precipitation comparable to, if not better than, that of the dynamical approach (i.e. the CFS).

Here, we present evidence that the use of autoregressive (AR) models with or without additional covariates can produce a model with reasonable skill in the prediction of Nepal's monsoon precipitation variation for up to 2 years. Such an accomplishment of AR models may result from the fact that the monsoon precipitation in Nepal is highly cyclical, being modulated by the so-called Pacific QDO in a phase-shifting manner (WG12; Wang *et al.*, 2011). The Pacific QDO describes the low frequency variability in the sea surface temperature (SST) over the central tropical Pacific (Allan, 2000) and was also found to modulate ENSO (White and Liu, 2008). WG12 found that the monsoon precipitation in Nepal is enhanced when southeasterly moisture fluxes originating from the Bay of Bengal are diverted towards the north and interact with the southern Himalayan foothills. The redirected moisture fluxes are modulated through global divergent circulation anomalies that are driven by, and evolving around, the lifecycle of the Pacific QDO. However, the modulation exhibits a phase shift of 2 years between the precipitation anomalies in Nepal and the extreme phases of the Pacific QDO in the central Pacific. Earlier, using a different lagged-phase relationship discovered to exist between the Pacific QDO and the elevation change of the Great Salt Lake in the western United States, Gillies *et al.* (2011) developed a set of coupled principal component-lagged regression models that is capable in forecasting the lake elevation out to 8 years with reasonable skill. That study was the theoretical basis for the development of a suitable AR model for the interannual prediction of Nepal's monsoon precipitation.

2. Data sources

The first step in the development of a suitable prediction model for all of Nepal was the application of a gridded precipitation data set, i.e. the Asian precipitation-highly resolved observational data integration towards evaluation of the water resources (APHRODITE) at a resolution of $0.5^{\circ} \times 0.5^{\circ}$; this dataset spans the period of 1951–2007 (Yatagai *et al.*, 2012). The APHRODITE project incorporated the full range of rain-gauge data from Nepal's Department of Hydrology and Meteorology and was able to construct the monsoon rainfall climatology that is superior to that revealed from either the Global Telecommunication System (GTS) network or the Global Precipitation Climatology Centre (GPCC) dataset (Yatagai *et al.*, 2012). Furthermore, the APHRODITE depiction of the monsoon rainfall in Nepal is in good agreement with that revealed from the Tropical Rainfall Measurement Mission (TRMM) satellite dataset. After the 1950s, the station network in Nepal expanded to over 80 stations that provided a relatively uniform coverage of the country (WG12, their Figure 2). For data prior to 1951, we adopted the gauge-based precipitation of the University of Delaware (UDEL) at a $0.5^{\circ} \times 0.5^{\circ}$ resolution (Legates and Willmott, 1990). The months of July and August represent the peak monsoon season with a dominant tropical forcing; thus, Nepal's monsoon was defined as the JA mean precipitation averaged from 52 grid points of the precipitation datasets within the boundary of Nepal. Following WG12, the APHRODITE and UDEL data were then combined to form a precipitation time series spanning the period 1931–2008. These two datasets exhibited a coherent quasi-decadal variability (WG12).

We also used observed SST data from the NOAA Extended Reconstructed SST (ERSST) Version 3b at a $2^{\circ} \times 2^{\circ}$ resolution (Smith *et al.*, 2008). The Pacific QDO was represented by the Nino-4 SST index based on the finding of Wang *et al.* (2011). Finally, seasonal hindcasts of the CFS version 2 (Saha *et al.*, 2010), obtained from the National Climatic Data Center at http://nomads.ncdc.noaa.gov/, were compared with the statistical model output.

3. Statistical models

The autoregressive (AR) process is one of the simplest and, at the same time, the most fundamental type of time series models that accommodates a large class of data which show serial correlations – a property that is effectively revealed by the autocorrelation function

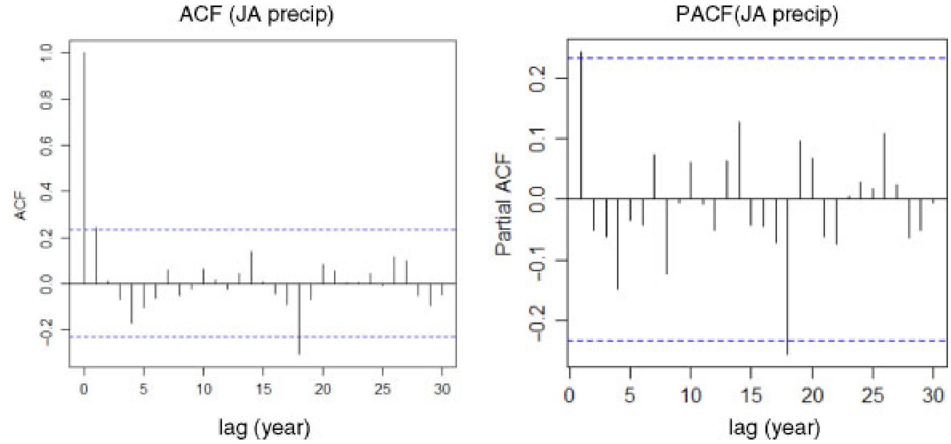

Figure 2. The autocorrelation function (ACF) and partial autocorrelation (PACF) of the JA mean precipitation in Nepal with the 95% confidence interval (blue dashed line). This figure is available in colour online at wileyonlinelibrary.com/journal/joc

(ACF) and the partial autocorrelation function (PACF; Brockwell and Davis, 1991). The ACF and PACF of Nepal's JA precipitation series (Figure 2) indicates that the series is significantly autocorrelated, particularly at lags 1 and 18 and this outcome was the provision for our initial justification that an AR process to model the data was appropriate. In this study, we developed three AR models: (1) using precipitation alone, (2) the combination of precipitation with the Nino-4 SST index (i.e. to assess the impact of the Pacific QDO on forecasting ability) and (3) the combination of precipitation with the effect of the global circulation of water vapor fluxes. In order to examine the extent to which AR modelling was apt in capturing the precipitation series, we compared the three models' prediction performance over a 50-year period (1959–2008) to that of (1) a persistent forecast (PF; i.e. next year's rainfall is the same as the present year) and (2) a climatological forecast (CF; i.e. every year's rainfall is the same as the long-term climatology). The comparison and evaluation of the three models are discussed in the next sections.

3.1. The precipitation-autoregressive (P-AR) model

Given our goal of forecasting the JA precipitation in Nepal, the first AR model is simply one of using the past precipitation data to forecast the future. Given the ACF and PACF of the JA precipitation, we first considered a full model that included lags from 1 to 18. The coefficients of the AR model were first determined using the final prediction error (FPE) criterion (Akaike, 1970). The model size was then reduced using a sparse coefficient estimation procedure first developed by Sun and Lin (2012) and later refined by Sang and Sun (2012), based upon the penalized conditional maximum likelihood (PCML). Specifically, Sang and Sun (2012) found that the smoothly clipped average deviation (SCAD) penalty performs better than other popularly used penalty functions (i.e. one which serves to guard against over-fitting) when it comes to sub-model selection and parameter estimation. More importantly, they demonstrated that the PCML-SCAD method not only consistently selects the optimal sub-model, but also estimates the selected sub-model parameters with a higher accuracy and efficiency. Therefore, using SCAD PCML, the AR model of choice was defined accordingly as the P-AR model:

$$\begin{aligned}P_t = {}& 0.1033P_{t-1} - 0.0185P_{t-4} - 0.1369P_{t-5} \\ & -0.1248P_{4-6} + 0.0370P_{t-7} - 0.0955P_{t-8} \\ & +0.0403P_{t-10} + 0.0230P_{t-11} - 0.0603P_{t-12} \\ & +0.0567P_{t-14} - 0.0372P_{t-15} - 0.1480P_{t-16} \\ & -0.1154P_{t-17} - 0.3965P_{t-18} + Z_t.\end{aligned}$$

Here, P_t denotes the centred (i.e. de-meaned or departure) JA mean precipitation at the present year while $t-n$ denotes the period for the previous n year. Z_t represents a white noise process with a zero mean and a variance equal to 902.42. As AR models are assumed to have a zero mean, this study centred all the datasets before fitting the model. In other words, we were dealing with precipitation anomalies which have a mean equal to zero. Figure 3(a) displays the model predicted value *versus* the observed centred JA precipitation. The R squared of the model was 0.487. In the 50-year forecasting period, the mean absolute error (MAE) was 0.94 mm day^{-1} and the root mean squared error (RMSE) was 1.16 mm day^{-1}. Both the P-AR MAE and the RMSE were smaller than those of either the PF or CF models as are summarized in Figure 4(a) and (b), respectively. The ACF and PACF of the model residuals fell within the 95% confidence interval (Figure 5(a)), indicating that the residuals were uncorrelated. Therefore one can assume that the majority of the data structure has been explained by the P-AR model.

3.2. The AR with covariate SST (ARX-SST) model

In general, an AR model that includes covariates (ARX) enables the model to take other factors into account that might be crucial to the AR time series (i.e. the

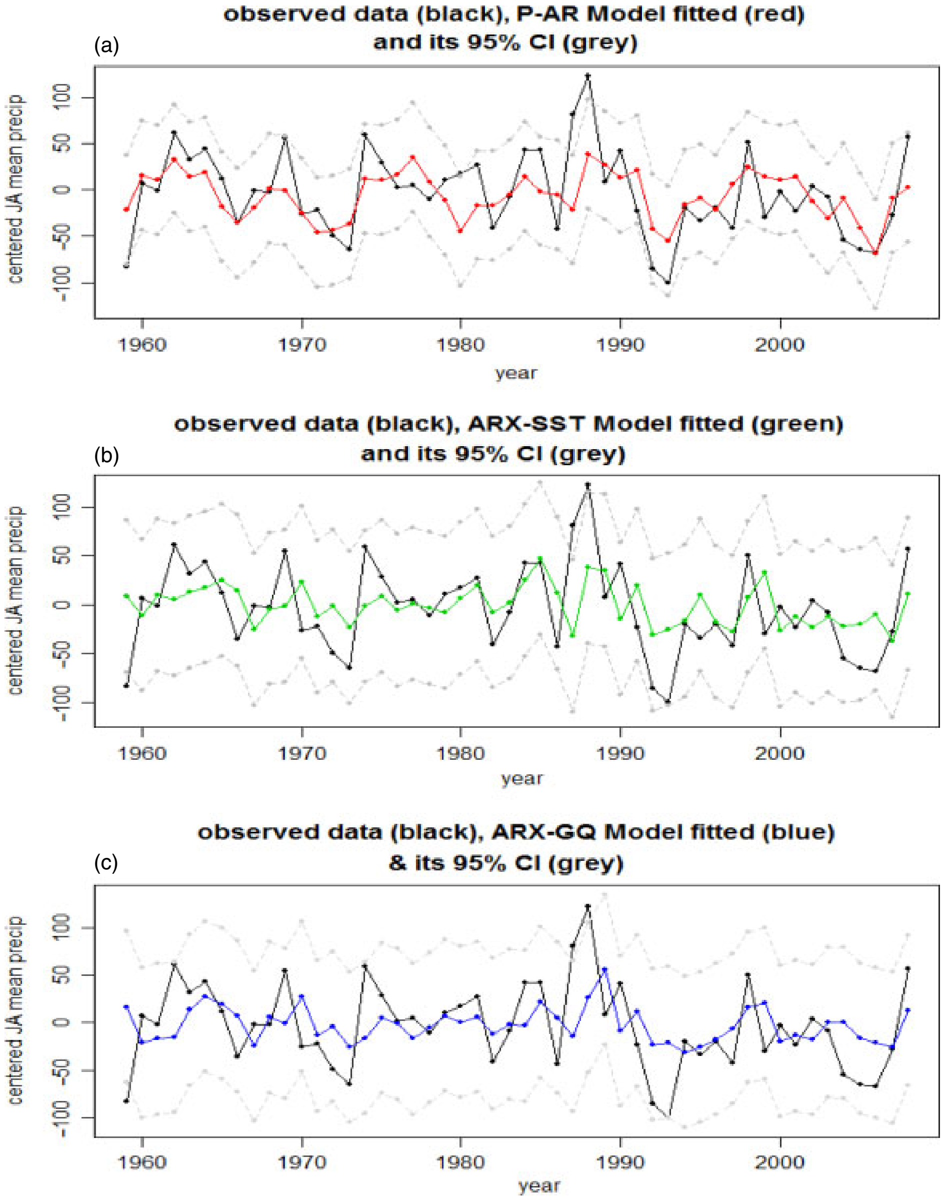

Figure 3. The model fitted data (colour line) *versus* observed data (black line) by the (a) P-AR, (b) ARX-SST and (c) ARX-GQ methods. The 95% confidence levels are given in light grey lines. This figure is available in colour online at wileyonlinelibrary.com/journal/joc

JA precipitation). Given prior evidence of the Pacific QDO's influence on precipitation in Nepal (WG12), we decided to examine if an AR model that included the influence from the Central Pacific SST (denoted as ARX-SST) might better forecast the precipitation variations. Following the procedures used in Wang *et al.* (2011), we adopted the Nino-4 SST anomalies to depict the Pacific QDO. Figure 6(a) shows the cross-correlation between the Nino-4 SST anomalies and the JA precipitation. Given that a 10- year oscillation is observed in the cross-correlation function, we included the 10-year lag of the Nino-4 SST anomalies into the model. The final model with Nino-4 SST anomalies (ARX-SST) was defined as:

$$P_t = -21.380SST_{t-10} + 0.335P_{t-1} - 0.114P_{t-2} + Z_t.$$

Here, as was the case for the P-AR model, we removed the mean of the Nino-4 SST anomalies series before fitting the model. SST_t denotes the centred Nino-4 SST. Z_t represents a white noise process with a zero mean and variance equal to 1575.11. The model fitted centred JA mean precipitation *versus* the observed are shown in Figure 3(b). In the 50-years forecasting period, both of the error rates were larger than those of P-AR model (Figure 4): the MAE was 1.06 mm day^{-1} while the RMSE was 1.35 mm day^{-1}. The ARX-SST ACF and PACF of the residuals fall within the 95% confidence interval (Figure 5(b)); this suggests that the residuals were uncorrelated at the significance level of 95% and the fitted model is a reasonable one.

3.3. AR with covariate water vapor fluxes circulation (ARX-GQ) model

WG12 reported that the global water vapor flux circulations, which respond strongly to Pacific QDO forcing, are

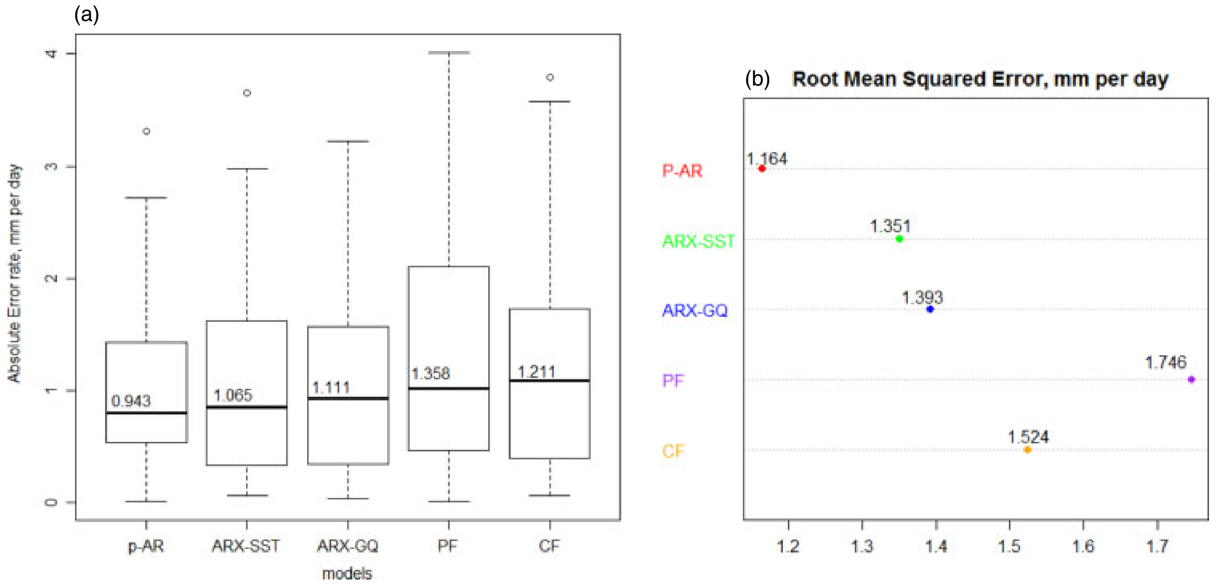

Figure 4. The (a) mean absolute error (MAE) and box plots and (b) root mean squared error (RMSE) for the five forecasting methods analyzed. This figure is available in colour online at wileyonlinelibrary.com/journal/joc

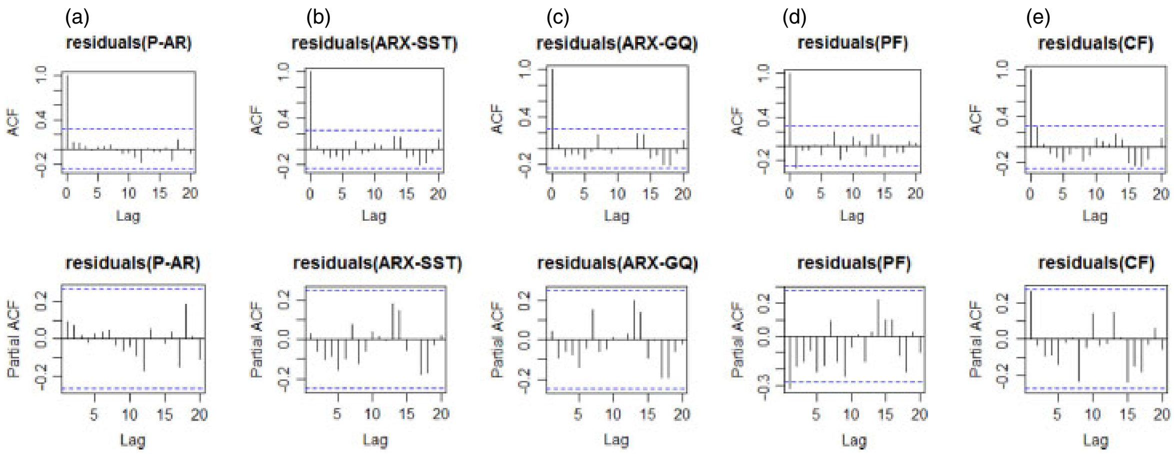

Figure 5. The ACF (top) and PACF (bottom) of the residuals for the (a) P-AR, (b) ARX-SST, (c) ARX-GQ, (d) PF and (e) CF methods with the 95% confidence interval (blue dashed line). This figure is available in colour online at wileyonlinelibrary.com/journal/joc

spectrally coherent with Nepal's monsoon precipitation and the coherence features a 1- to 2-year lag in the Nino-4 SST. Physically speaking, when the monsoon precipitation peaks in Nepal, southerly water vapor fluxes driven by the Pacific QDO circulations form substantial convergence over the Indo-Gangetic Plain 1–2 years prior to a cold-phase of the Pacific QDO. Conversely, northerly water vapor fluxes and divergence accompany dry summers 1–2 years prior to a warm-phase. Such regional circulation anomalies are embedded in an eastward displacement of the global divergent circulation anomalies associated with the Pacific QDO's lifecycle (Figure 8 of WG12). As a result, we decided to include the influence of the global water vapor flux circulation (GQ) in an ARX model. In WG12 the GQ index was computed based upon the leading principal component (PC) of the vertically integrated water vapor fluxes for the JA period; this was derived from the Twentieth Century Reanalysis (20CR) Version 2 (Compo *et al.*, 2011). The GQ index included in the ARX model was obtained directly from WG12.

After several trials we selected the following model with the lowest RMSE:

$$P_t = -11.704GQ_{t-7} + 6.179GQ_{t-8} + 4.624GQ_{t-9} + 24.618GQ_{t-10} + 0.326P_{t-1} + Z_t.$$

Since the unit length of the GQ index was one (owing to its origin from the normalized PC analysis), no further centring was needed. GQ_t here represents the GQ index at the present year. Z_t represents a white noise process with a zero mean and variance equal to 1646.11. In the 50-year forecasting period, the ARX-GQ model's MAE was 1.11 mm day^{-1}, while the RMSE was 1.39 mm day^{-1} (Figure 4) – these were the largest among all three AR models. The fact that the model residuals were not autocorrelated (Figure 5(c)) suggests that ARX-GQ remains a

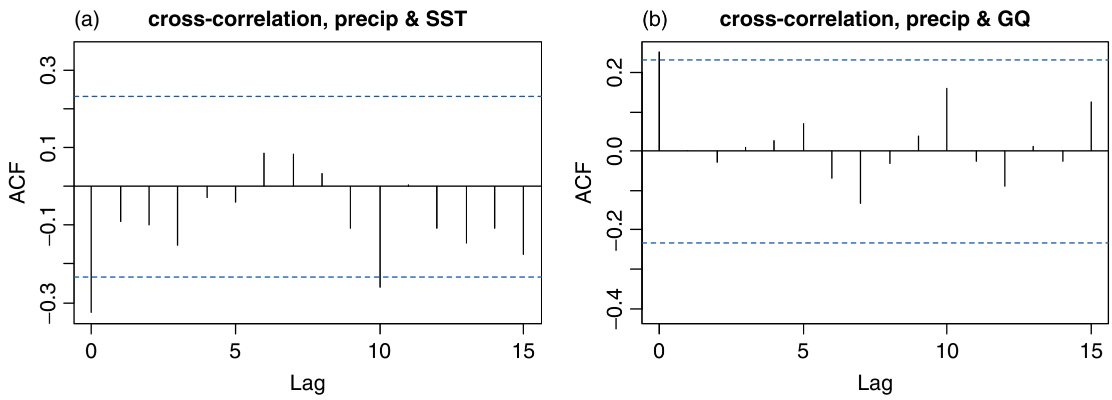

Figure 6. The cross-correlation function (a) between JA precipitation and SST and (b) between JA precipitation and GQ. This figure is available in colour online at wileyonlinelibrary.com/journal/joc

reasonable model. The model fitted centred JA mean precipitation *versus* the observed are shown in Figure 3(c).

3.4. The persistent and climatological forecasts

As shown in Figure 4, the PF model's MAE and RMSE were 1.358 and 1.746 mm day^{-1}, respectively, and were greater than the three AR models. The ACF and PACF plots (Figure 5(d)) of the residuals showed that the residuals were autocorrelated at the 95% significance level meaning that there exists structure in the data that is unexplained by the model. In regards to the CF model, the 20-year mean was used to forecast for the present year. The CF's MAE and RMSE were 1.211 and 1.524 mm day^{-1}, respectively, somewhat less than those of the PF. The CF model does not capture any of the monsoon variability (Figure 5(e)) hence it is not comparable.

4. Model inter-comparison and evaluation

To summarize, the performance of all three AR models surpassed that of both the PF and CF. Using the P-AR model's MAE value (i.e. 0.943 mm) as a point of reference, both the PF's and CF's MAE were larger-on the order of 45% and 29%, respectively. Correspondingly, the ARX-SST's and ARX-GQ's MAE were on the order of 14% and 18% greater that the P-AR's. Likewise, applying the same criterion as for the RMSE, the PF's and CF's RMSE were 51% and 31% greater, while ARX-SST's and ARX-GQ's were, respectively, 16% and 20% larger. Moreover, the PF is, by definition, only able to predict 1-year ahead.

As all three AR models are theoretically able to predict more than 1 year (i.e. predict the forthcoming year and the years after), we decided to further examine the AR models' performance through a 2-year prediction; this being based on the fact that Nepal's monsoon precipitation and the Pacific QDO lifecycle exhibits a 1- to 2-year phase lag (WG12). Thus, we conducted ten cross-validations in which each involved a 2-year forecast. All model coefficient βs were kept the same, but the forecasting

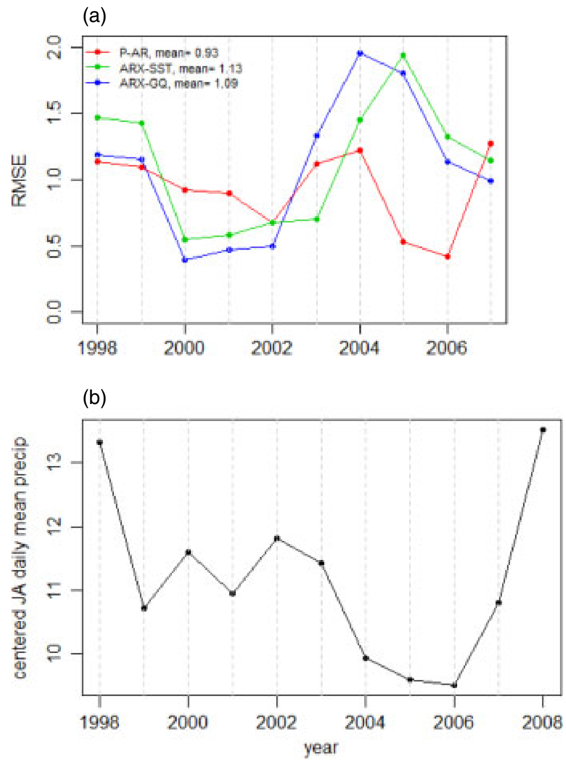

Figure 7. (a) Ten two-step ahead prediction RMSE's of the three AR models *versus* (b) the observed JA mean precipitation amounts. This figure is available in colour online at wileyonlinelibrary.com/journal/joc

errors were estimated exclusively over the 2-year forecasting horizon. The initial forecasting period was set at 1998–1999. The model was fitted up to 1997 and the predicted values were then used to forecast the precipitation amounts in 1998 and 1999. The second forecasting period was 1999–2000 with subsequent years being incremented accordingly; hence the final forecasting period was 2007–2008. The root mean squared errors (RMSE) were subsequently computed. The results are shown in Figure 7(a). The P-AR model as the one with the smallest RMSE range (i.e. 0.42–1.28 mm) is also noteworthy in that it has the lowest error mean in both the 1- and 2-year predictions. The performances of the ARX models differ

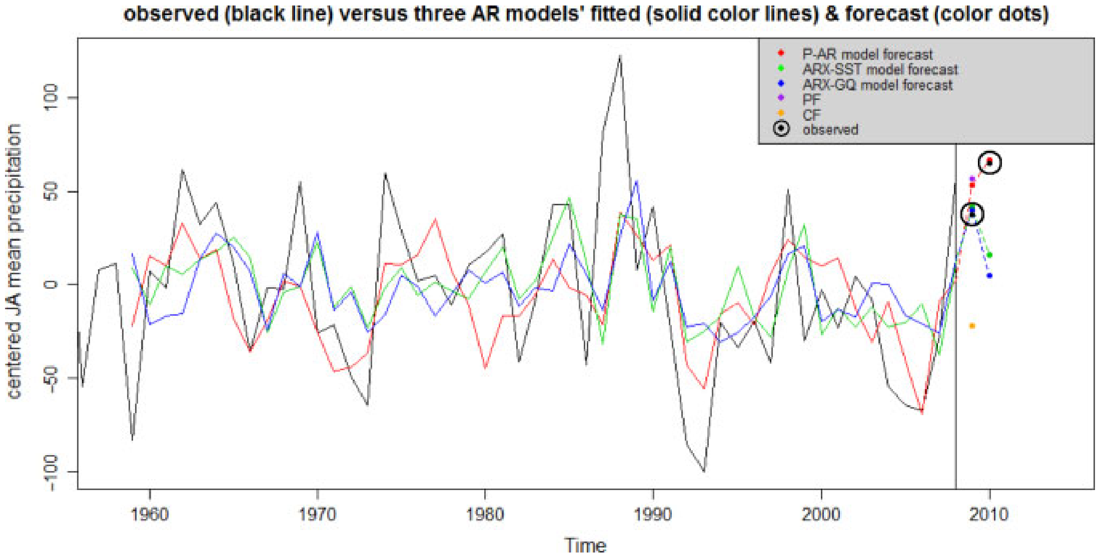

Figure 8. Comparison of the three AR models' fitted values (solid colour lines) and the 2009–2010 JA precipitation forecast (dash colour lines). PF (purple dot) and CF (orange dot) gives 2009 forecast only. Values of the observed for 2009 and 2010 (black dots) were derived from an independent dataset PREC/L (see text). This figure is available in colour online at wileyonlinelibrary.com/journal/joc

from that of the P-AR model over the 10-year period. The ARX-SST model's error ranged from 0.40 to 1.9 mm with a mean of 1.13, while ARX-GQ model's error ranged from 0.40 to 1.96 mm, with a mean of 1.09. The performance of ARX-GQ was observed to be consistently better than that of ARX-SST model except when the JA precipitation dropped substantially in 2003 and 2004 (Figure 7(b)). In this regard, the P-AR model outperforms the ARX models as it captures the variability better.

Finally, to investigate the extent to which the AR models provide any added value to the seasonal prediction, we compared each of them with the CFS's precipitation forecast. The CFS precipitation forecast was derived from the May initial conditions using a 10 member ensemble (i.e. a 2-month lead time for the July to August monsoon season). In order that the AR model results would be compatible and comparable, the CFS precipitation forecast output was interpolated bilinearly onto the APHRODITE grid at a $0.5^\circ \times 0.5^\circ$ spatial resolution and, the precipitation amounts were averaged within the geographical boundary of Nepal. The July to August mean errors of each year were removed. We then computed the RMSE between the observed precipitation and the derived CFS precipitation fields (results not shown). The results (not shown) revealed that, even with the monthly mean errors removed, the CFS produced a precipitation RMSE of more than 2.5 times larger in comparison to each of the three AR models. Such a large bias is consistent with the evaluation undertaken by Pattanaik and Kumar (2010; their Figure 8) that showed that the CFS consistently over-predicts precipitation in Nepal. However, the result here should be interpreted with caution, because the 9-month seasonal prediction of CFS does not cover the 2-year forecast of the statistical models; hence the CFS comparison made here should be perceived as suggestive rather than conclusive.

5. Conclusions and discussion

In this paper, we compared three different AR models in the prediction of summer precipitation amounts for Nepal and evaluated their performances by two standard statistical criteria: MAE and RMSE. In addition, we compared the AR models with (1) a persistent method and (2) a climatology method as these afforded us a heuristic and non-statistical use of the historical data as predictors. The three AR models were categorized as follows: the first used the antecedent JA precipitation data alone, while the other two included Nino-4 SST anomalies and the GQ index respectively. The results indicated that the P-AR model maintained a stable forecasting performance, especially in periods of extreme precipitation anomalies. Both ARX-SST and ARX-GQ models performed well when the observed precipitation anomalies are non-extreme; this suggests that both GQ and Nino-4 SST anomalies could serve as an indicator for mild precipitation changes, but large precipitation changes tend to be cyclic in nature and therefore were better captured by the auto-regressive method.

When using all the five methods to forecast Nepal's JA precipitation in 2009, the results shown in Figure 8 indicate that PF and CF had counter predictions – the PF predicted an above-normal amount while the CF predicted a below-normal amount. For the 2009-2010 summers, all three AR models predicted above-normal amounts. Verification of the forecasted 2009–2010 precipitation amounts was made with the NOAA's Precipitation Reconstruction over Land (PREC/L) with a $1^\circ \times 1^\circ$ resolution (Chen *et al.*, 2002); this provided, in a loosely comparative sense, how the precipitation amounts have fluctuated in Nepal outside the APHRODITE data period. What emerged from the comparison with NOAA's PREC/L (Figure 8) is that all three AR models predicted the mild anomaly in 2009 (a slightly strong monsoon) while the

P-AR model was the only one to successfully forecast the increased amount in 2010 (a strong monsoon).

In conclusion, it appears that it is possible to predict the monsoon precipitation in Nepal with reasonable accuracy out to 2 years through the application of simple time series models. Perhaps of more significance, however, is that the given statistical predictions outperform current operational models (in this case the CFS), which has a limited skill up to 2 months. The statistical models presented here, while they may be considered simplistic, are in effect a viable alternative, if not a valuable approach towards predicting Nepal's monsoon variability. Due to the different performance of P-AR and ARXs in different amplitude zones of the anomalous precipitation, the combined usage of P-AR and ARX models may achieve a better skill, provided that the Pacific QDO tendency is forecasted. In essence then, they should be taken into consideration as an intermediate solution for climate precipitation forecasting in lieu of the time when climate dynamical models will be more reliable.

Acknowledgements

This study was supported by the United States Agency for International Development through Grant no. EEM-A-00-10-00001, and by the Utah State University Agricultural Experiment Station.

References

Achuthavarier D, Krishnamurthy V, Kirtman BP, Huang B. 2012. Role of the Indian Ocean in the ENSO-Indian summer monsoon teleconnection in the NCEP Climate Forecast System. *Journal of Climate* **25**: 2490–2508. DOI: 10.1175/JCLI-D-11-00111.1.

Akaike H. 1970. Statistical predictor identification. *Annals of the Institute of Statistical Mathematics* **22**: 203–217.

Allan RJ. 2000. ENSO and climatic variability in the last 150 years. In *El Niño and the Southern Oscillation: Multiscale Variability, Global and Regional Impacts*, Diaz HF, Markgraf V (eds). Cambridge University Press: Cambridge, UK; 3–56.

Brockwell P, Davis R. 1991. *Time Series: Theory and Methods*, 2nd edn. Springer-Verlag: New York.

Chen M, Xie P, Janowiak JE, Arkin PA. 2002. Global land precipitation: a 50-yr monthly analysis based on gauge observations. *Journal of Hydrometeorology* **3**: 249–266.

Compo GP, Whitaker JS, Sardeshmukh PD, Matsui N, Allan RJ, Yin X, Gleason BE, Vose RS, Rutledge G, Bessemoulin P, Brönnimann S, Brunet M, Crouthamel RI, Grant AN, Groisman PY, Jones PD, Kruk MC, Kruger AC, Marshall GJ, Maugeri M, Mok HY, Nordli Ø, Ross TF, Trigo RM, Wang XL, Woodruff SD, Worley SJ. 2011. The twentieth century reanalysis project. *Quartely Journal of the Royal Meteorological Society* **137**: 1–28.

DelSole T, Shukla J. 2002. Linear prediction of Indian monsoon rainfall*. *Journal of Climate* **15**: 3645–3658.

Gillies RR, Chung O-Y, Wang S-Y, Kokoszka P. 2011. Incorporation of Pacific SSTs in a time series model towards a longer-term forecast for the Great Salt Lake elevation. *Journal of Hydrometeorology* **12**: 474–480.

Legates DR, Willmott CJ. 1990. Mean seasonal and spatial variability in gauge-corrected, global precipitation. *International Journal of Climatology* **10**: 111–127.

Navone HD, Ceccatto HA. 1994. Predicting Indian monsoon rainfall: a neural network approach. *Climate Dynamics* **10**: 305–312.

Pattanaik DR, Kumar A. 2010. Prediction of summer monsoon rainfall over India using the NCEP climate forecast system. *Climate Dynamics* **34**(4): 557–572.

Pattanaik DR, Kumar A, Tyagi A. 2010. Development of empirical-dynamical hybrid forecasts for the Indian monsoon rainfall using the NCEP Climate Forecast System. Synoptic Meteorology No: 11/2010, IMD Met Monograph, Indian Meteorological Department. http://www.imd.gov.in/section/nhac/dynamic/Report_AISMR.pdf.

Saha S, Nadiga S, Thiaw C, Wang J, Wang W, Zhang Q, Van den Dool HM, Pan H-L, Moorthi S, Behringer D, Stokes D, Peña M, Lord S, White G, Ebisuzaki W, Peng P, Xie P. 2006. The NCEP climate forecast system. *Journal of Climate* **19**: 3483–3517.

Saha S, Moorthi S, Pan H-L, Wu X, Wang J, Nadiga S, Tripp P, Kistler R, Woollen J, Behringer D, Liu H, Stokes D, Grumbine R, Gayno G, Wang J, Hou Y-T, Chuang H-Y, Juang H-MH, Sela J, Iredell M, Treadon R, Kleist D, Van Delst P, Keyser D, Derber J, Ek M, Meng J, Wei H, Yang R, Lord S, Van Den Dool H, Kumar A, Wang W, Long C, Chelliah M, Xue Y, Huang B, Schemm J-K, Ebisuzaki W, Lin R, Xie P, Chen M, Zhou S, Higgins W, Zou C-Z, Liu Q, Chen Y, Han Y, Cucurull L, Reynolds RW, Rutledge G, Goldberg M. 2010. The NCEP climate forecast system reanalysis. *Bulletin of the American Meteorological Society* **91**: 1015–1057.

Sang H, Sun Y. 2012. Simultaneous sparse model selection and coefficient estimation for heavy-tailed autoregressive process. *Statistics* http://arxiv.org/abs/1112.2682.

Shrestha AB, Karmacharya J. 2006. Performance of nested RegCM3 climate model in central Himalayan range (A33A-05). American Geophysical Union Fall Meeting, San Francisco, CA, December 2007.

Shrestha AB, Wake CP, Dibb JE, Mayewski PA. 2000. Precipitation fluctuations in the Nepal Himalaya and its vicinity and relationship with some large scale climatological parameters. *International Journal of Climatology* **20**: 317–327.

Shukla J, Mooley DA. 1987. Empirical prediction of the summer monsoon rainfall over India. *Monthly Weather Review* **115**: 695–703.

Smith TM, Reynolds RW, Peterson TC, Lawrimore J. 2008. Improvements to NOAA's historical merged land–ocean surface temperature analysis (1880–2006). *Journal of Climate* **21**: 2283–2296.

Sun Y, Lin X. 2012. Regularization for stationary multivariate time series. *Quantitative Finance* **12**(4): 573–586.

Wang S-Y, Gillies RR. 2012. Influence of the Pacific quasi-decadal oscillation on the monsoon precipitation in Nepal. *Climate Dynamics*. DOI: 10.1007/s00382-012-1376-2.

Wang S-Y, Gillies RR, Hipps LE, Jin J. 2011. A transition-phase teleconnection of the Pacific quasi-decadal oscillation. *Climate Dynamics* **36**: 681–693.

White WB, Liu Z. 2008. Resonant excitation of the quasi-decadal oscillation by the 11-year signal in the Sun's irradiance. *Journal of Geophysical Research* **113**: C01002.

Yang S, Zhang Z, Kousky VE, Higgins RW, Yoo S-H, Liang J, Fan Y. 2008. Simulations and seasonal prediction of the Asian summer monsoon in the NCEP climate forecast system. *Journal of Climate* **21**: 3755–3775.

Yatagai A, Kamiguchi K, Arakawa O, Hamada A, Yasutomi N, Kitoh A. 2012. APHRODITE: Constructing a long-term daily gridded precipitation dataset for Asia based on a dense network of rain gauges. *Bulletin of the American Meteorological Society* **93**: 1401–1415.

