Predicting Central Asian river flows from regional precipitation and wind patterns during the preceding cold season

MATHEW BARLOW[1] & MICHAEL TIPPETT[2]

1 *AER Inc., 131 Hartwell Avenue, Lexington, Massachusetts 02421-3126, USA*
mbarlow@aer.com

2 *IRI of Columbia University, Lamont-Doherty Earth Observatory, PO Box 1000, Palisades, New York 10964-8000, USA*

Abstract Snowmelt is an important driver of regional river flow, and warm season (April–August) river flows are linked to the previous cold season's (November–March) precipitation. Canonical correlation analysis (CCA) is applied between the cold season regional climate (November–March zonal winds and precipitation) and the subsequent warm season river discharge (April–August station river flows) for 1950–1985. The NCEP/NCAR re-analysis is used for precipitation and winds, and 24 stations in the mountains of eastern Uzbekistan and Tajikistan are used for river flow data. The extracted cold season precipitation and wind patterns have regional scales and are adequately captured by the NCEP/NCAR reanalysis. Average cross-validated skill correlation is 0.43 for the river flows, with 10 stations correlated greater than 0.5. As the re-analysis data are updated in real time, this scheme can make operational forecasts. The regional variability is also related to tropical Pacific sea-surface temperatures, which may enable forecasts at longer leads.

Key words Central Asia; river flow; seasonal forecasting; snowmelt; streamflow

INTRODUCTION

Central Asia is a semiarid region, with precipitation primarily occurring during the cold season (November–April), and falling as snow in the high mountains of the region. As a result of this seasonality, snowmelt is a primary driver of the regional river flows which tend to peak between April and August. There is considerable year-to-year variability in regional precipitation, including a recent severe drought (Barlow *et al.*, 2002), and due to water scarcity, infrastructure problems, and large irrigation demands, there is high societal vulnerability to the variability. Global warming is also a concern, as temperature increases have been associated with notable decreases in the local glaciers (e.g. Aizen *et al.*, 1997) and circulation changes associated with the warming trend in tropical sea-surface temperatures (SSTs) appear to be related to the recent drought (Barlow *et al.*, 2002; Hoerling & Kumar, 2003). The regional precipitation is also related to large-scale, predictable climate variability (Barlow *et al.*, 2002; Tippett *et al.*, 2003). These teleconnections are used as a basis for operational predictions of cold season precipitation in the region (iri.columbia.edu/climate/forecast/cswasia/index.html), with demonstrated skill (Tippett *et al.*, 2004).

The seasonal lag between the accumulation of snow and its subsequent melting suggests an additional forecasting tool for warm season hydrological variables. An

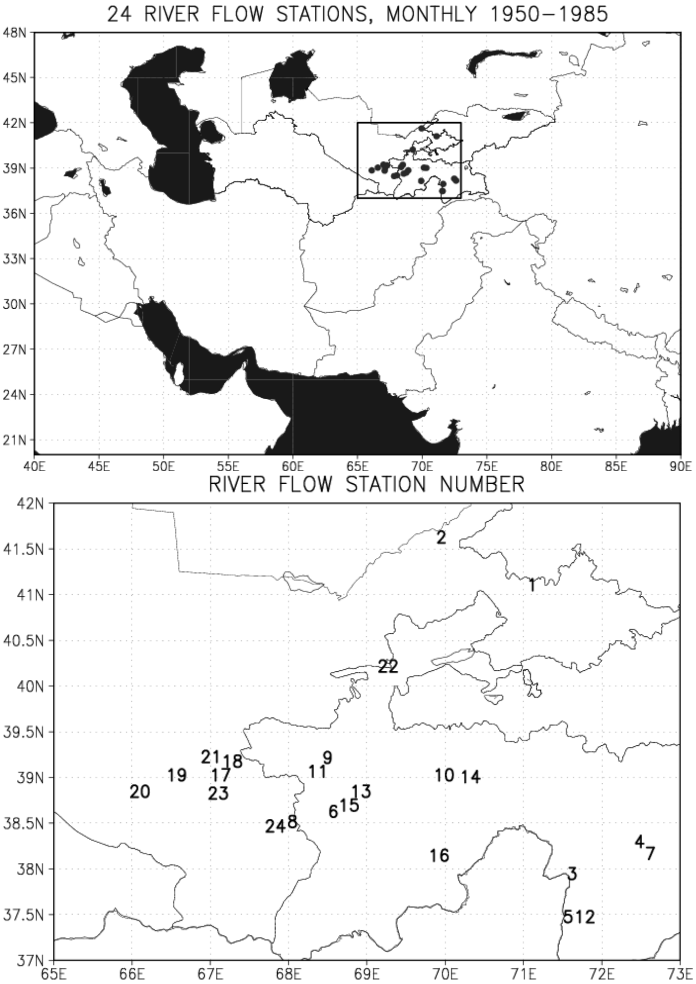

Fig. 1 Location of the river flow stations. More information about the stations is given in Table 1.

accurate estimate of snow amount just prior to the beginning of the spring melt would provide a prediction of the subsequent river flows. Unfortunately, a sufficiently accurate estimate of snow or, more specifically, snow water equivalent, is very difficult to obtain, particularly in real time. However, as most precipitation in the mountains falls as snow, accumulated precipitation may provide a sufficient proxy. This approach has been explored by Schär *et al*. (2004), who looked at prediction of the Syr Darya and Amu Darya basin discharges from accumulated cold season precipitation averaged over the basin drainage area. Since well-sampled precipitation observations are also difficult to obtain for the region, Schär *et al.* used the model-based precipitation product from the ECMWF reanalysis, with good results for the Syr Darya. They note that, while the ECMWF reanalysis also suffers from the sparseness of local observations, storms arriving in the region are very well sampled just upstream in Europe and the local distribution of precipitation is strongly controlled by the orography. These factors are included in dynamically-based re-analyses, such as those of the ECMWF and NCEP/NCAR, and can provide information even in regions with few local observations.

Table 1 CCA results for river flow station data, shown as correlation. Cross-validated correlations are also shown. For station numbers refer to Fig. 1. An asterisk denotes statistical significance at the 99% or higher level, based on Monte Carlo analysis. Note that the negative correlations of river flow shown here correspond to negative precipitation anomalies in Fig. 2; that is, the river flow and precipitation anomalies have the same sign, as expected.

Station name	River name	Drainage area (km^2)	Elevation (m)	Apr–Aug fraction of annual river flow	CCA correlation	Cross-validated correlation
1. Gava	Gavasay	657	1063	0.83	*–0.85	*–0.61
2. Khujakent	Ugam	869	741	0.72	*–0.91	*–0.59
3. Shuchand	Murgab	24 700	90	0.58	*–0.48	–0.08
4. Barchadiv	Murgab	16 700	2510	0.40	–0.26	+0.05
5. Khorog	Gunt	13 700	2070	0.73	*–0.69	*–0.43
6. Alibegi	Khanaka	362	1004	0.75	*–0.61	*–0.44
7. Mouth of	Lyangar	335	3165	0.63	–0.25	+0.07
8. Dashnabad	Obizarang	330	768	0.75	*–0.58	*–0.38
9. Pinen	Pasrut	340	1770	0.66	*–0.85	*–0.54
10. Post-Sangikar	Sangikar	291	1279	0.77	*–0.87	*–0.71
11. Mouth	Sarytag	537	2195	0.81	*–0.76	*–0.49
12. Khabost	Shakhdara	4 180	4096	0.74	*–0.76	*–0.39
13. Takob	Tuykutal	140	1473	0.74	*–0.63	*–0.49
14. Garm	Vakhsh	20 000	1300	0.69	*–0.65	*–0.40
15. Dagana	Varzob	1 270	969	0.77	*–0.88	*–0.55
16. Karboztonak	Yakhsu	1 440	982	0.75	*–0.68	–0.30
17. Khozar-Nova	Akdarya	845	971	0.76	*–0.86	*–0.57
18. Dzhauz	Dzhindydarya	152	6	0.54	*–0.61	*–0.46
19. Chirakchi	Kashkadarya	4 970	510	0.54	*–0.79	*–0.52
20. Karatikon	Kashkadarya	7 900	411	0.66	*–0.86	*–0.58
21. Varganza	Kashkadarya	511	818	0.57	*–0.83	*–0.52
22. Bekabad	Syr Darya	142 000	292	0.83	*–0.61	*–0.43
23. Tatar	Yakkabag darya	504	1183	0.82	*–0.83	*–0.55
24. Dupuli	Zeravshan	10 200	1041	0.74	*–0.75	*–0.48

The importance of regional-scale climate patterns in the precipitation variability, as shown by Barlow *et al*. (2002) and Tippett *et al*. (2003), suggests a complementary approach: using the relationship between the river flows and the regional-scale precipitation and wind fields (as opposed to only considering the precipitation local to the basin). Consideration of the regional-scale and multiple variables includes more data, which may further alleviate some of the problems with sparsely observed local data. Accordingly, we apply the canonical correlation analysis (CCA) approach as in Tippett *et al.* (2003) to extract the joint patterns of variability between regional winds and precipitation during November–March (Nov–Mar) and river discharge station data during April–August (Apr–Aug). Because of the choice of non-overlapping seasons, the CCA can also be used to make forecasts: once the Nov–Mar data are available (generally by 5 April), the Apr–Aug average river flows—the main flow season—can be forecast.

Here we focus on eastern Uzbekistan and Tajikistan, where there are 24 river discharge stations (Fig. 1) with near-continuous monthly reports from 1950–1985. Based on a number of considerations, the data quality appears adequate for the current analysis; this is discussed in the next section. The NCEP/NCAR re-analysis is used for both winds and precipitation; using observed precipitation yields better results for earlier periods when regional coverage was more comprehensive but the re-analysis has the advantage of being operationally updated in a consistent fashion.

DATA

The river flow data was obtained from the National Center for Atmospheric Research (NCAR) dataset ds553.2. Twenty-four stations are considered, with less than 7% missing monthly data for any station (11 stations have no missing data) for Apr–Aug, 1950–1985. The station names, rivers, basin areas, and elevations are given in Table 1 and the locations are shown in Fig. 1. Observational precipitation was obtained from the extended New *et al*. (2000) gridded data. Model-based precipitation and upper-level zonal wind were obtained from the NCEP/NCAR reanalysis (Kalnay *et al*., 1996).

The general fidelity of the re-analysis variables in the region has been shown by Tippett *et al*. (2003). The quality of the river flow data has been evaluated with respect to physical consistency and the results of independent analysis. The primary relationship present in the river flow data, as shown in the next section, is with the cold season precipitation, reflecting the fundamental physical process of snowmelt. The patterns of the cold season precipitation extracted by CCA with the river flow data are similar to previous analysis of regional precipitation (Tippett *et al*., 2003). The importance of cold season precipitation to basin discharge in the region has been demonstrated by Schär *et al*. (2004) in a study that included post-1985 data and explicitly corrected for human influence. The physical consistency of the derived relationships and the agreement with independent analyses suggest that the data quality of the flow rates used here is acceptable.

RESULTS

CCA is applied between gridded regional Nov–Mar climate (200hPa zonal wind and precipitation) and the subsequent Apr–Aug river flows (24 river discharge stations) for 1950–1985. The river flow data are normalized by their standard deviation, so that the results will not be dominated by a few large rivers. The resulting patterns in zonal wind and precipitation are shown in Fig. 2 and the resulting pattern in river flows is given in terms of station correlations in Table 1. Note the regional scales of the climate variables and the very high correlations obtained for the majority of the stations, despite wide ranges in the size of the drainage basins and elevations of the rivers. The highest correlations are spread throughout the region and are not simply the most downstream stations with the largest drainage basins. The two easternmost stations in the region (stations 4 and 7 in Fig. 1(b)) do have low correlations. These stations have later peak flows than the others and need to be considered separately. To verify that the re-analysis precipitation, which is a product of the underlying atmospheric model and not directly constrained by observations, is contributing to a physical signal, the CCA time series was correlated to gridded observed precipitation (not shown); there is a good correspondence.

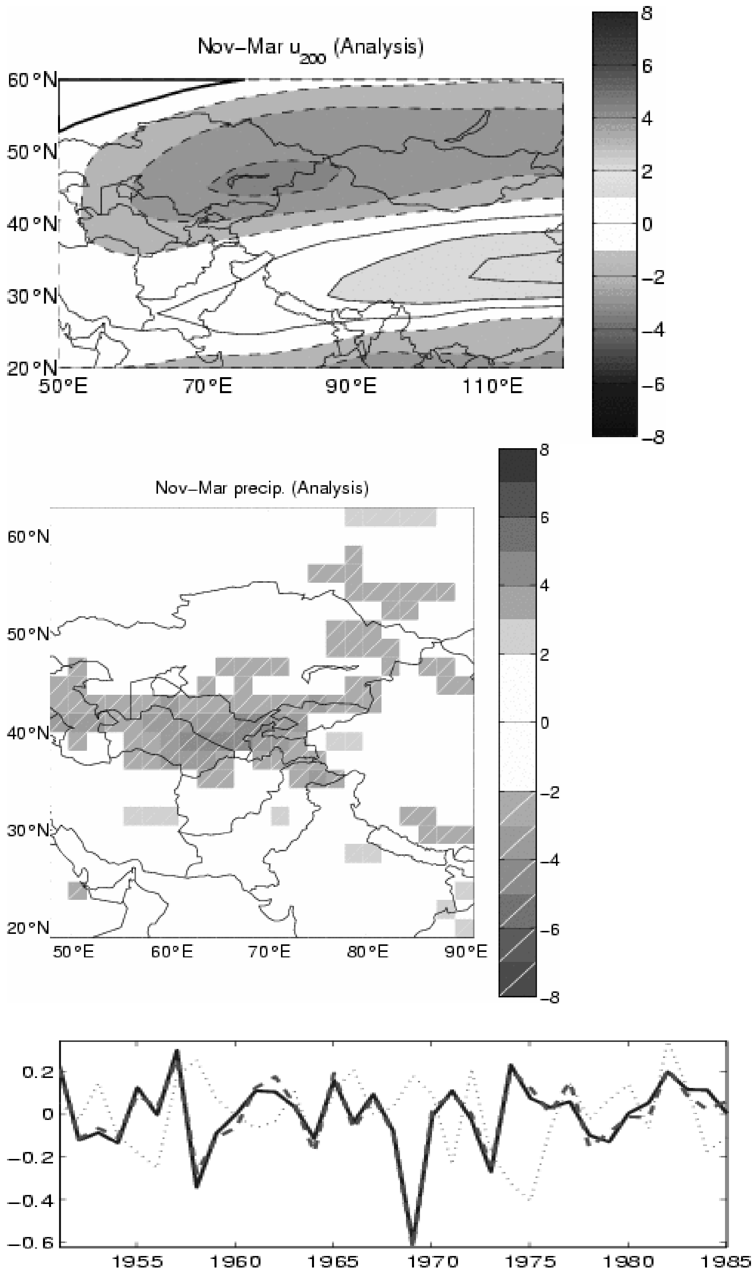

Fig. 2 CCA results: (a) re-analysis 200 hPa zonal wind, (b) re-analysis model precipitation, and (c) time series. In (c) the NINO4 index is also shown, indicated by a thin dotted line.

The variability captured by the CCA reflects the dominant local pattern of variability. Correlations to a simple box average of observed precipitation (65°–73°E, 37°–42°N) yield patterns very similar to those seen in Fig. 2, as do correlations to an average of all the normalized river flows.

There is considerable similarity between the patterns obtained in the current CCA analysis and those obtained from CCA between cold season winds and cold season precipitation in Tippett *et al*. (2003), although the maximum precipitation anomalies in the current analysis are located about 2° further northward and have greater westward extent. Decreased precipitation is associated with a reduction in jet-level winds, and has the largest changes along the windward slopes of the Pamir and Tien Shan ranges. Correlations to Pacific SSTs (not shown) are similar to those in Barlow *et al.* (2002) and Tippett *et al*. (2003), with a pattern similar to El Niño but with more emphasis on the central and western Pacific. The NINO4 time series, which is an average of tropical SSTs in the central Pacific, is also shown in Fig. 2(c) and is correlated with the CCA time series at –0.59.

The CCA may also be used for forecasting by projecting the precipitation and wind patterns onto the observed precipitation and wind anomalies for a given cold season. This magnitude multiplied by the pattern in river flows yields a river flow forecast for the subsequent warm season. To provide a robust estimate of forecast skill, a cross-validation is performed: the patterns are calculated leaving one year out of the data, then the prediction is made for that year; as all years are sequentially left out and forecast for, a skill is accumulated. The cross-validated skill—the skill likely to be obtained with operational forecasting—is shown in the last column of Table 1. Note that the average cross-validated skill is still above 0.5 for 10 of the 24 stations, and one station has a cross-validated correlation of 0.71 (50% of the variance). As NCEP/NCAR re-analysis data are updated operationally (about 5 days past the end of the month), the current approach can be used for operational forecasting.

SUMMARY AND DISCUSSION

The CCA shows that there is a close relationship between warm season river flows and the regional climate during the previous cold season—and that the operationally available NCEP/NCAR data are sufficient to capture this relationship. While this approach is not optimized for a particular station, it shows generally high correlations for most of the stations, despite the wide range of basin sizes, elevations, and upstream/downstream relationships present. This suggests that the results may be generally applicable to other river flow stations under the influence of the regional precipitation pattern. There are several additional stations in the region, which had fewer reports and so were not included in the CCA analysis but provide a useful validation; some of these stations also are highly correlated (discussed further below).

This study is limited by the lack of recent data, particularly given the dissolution of the Soviet Union, which changed water usage in the region. However, the regional patterns of cold season precipitation variability remain active and stable in the post-1985 periods (Tippett *et al*., 2003), and basin-averaged cold season precipitation continues to be a good predictor of local river flows in the recent period (Schär *et al*., 2004), so the CCA approach may be expected to have continued applicability. Indeed, preliminary comparisons with river flow data from the post-1985 period suggest that the regional CCA approach remains quite successful; we are in the process of collecting more data and validating this.

The CCA can also be employed to make predictions, and cross-validated correlations remain considerable, suggesting the technique, which is already in an operationally-useable format, may have practical utility. This regionally-based methodology appears to be a complementary approach with respect to the basin-accumulated precipitation used by Schär *et al.* (2004). The latter produced high correlation for the Syr Darya but low correlation for the Amu Darya. The Syr Darya at Bekabad (40.22°N, 69.27°E) is included in the present CCA. For 1950–1985, the correlation is 0.61 and the cross-validated correlation is 0.52; these are moderate correlations but not as high as achieved using basin-accumulated precipitation. The Amu Darya is not included in the present CCA, although some of its tributaries are, but it is present in the NCAR data with 25 years of data (1950–1974) at the Chatly (42.28°N, 59.7°E) station. Correlation with the CCA analysis is 0.83; cross-validated correlation is 0.73, considerably higher than is obtained using basin accumulations.

It appears that due to the precipitation seasonality favouring the cold season and the importance of snowmelt, regional river flows have good predictability even from the available data, and that advantage can be gained from both local and regional data. Moreover, the connection to tropical SSTs suggests the potential for even longer lead predictions, perhaps starting in the previous fall for the subsequent summer. In other areas of the world, such an approach has proven useful (Berri & Flamenco, 1999). Future research is planned to examine the seasonal predictability of central and southwest Asian river flows based on using different variables and combination of variables, as well as different domain sizes.

Acknowledgements This research was supported by the National Science Foundation Water Cycle programme under grant ATM-0233563.

REFERENCES

Aizen, V., Aizen, E., Melack, M. & Dozier, J. (1997) Climate and hydrologic changes in the Tien Shan, central Asia. *J. Climate* **10**, 1393–1404.

Barlow, M., Cullen, H. & Lyon, B. (2002) Drought in Central and Southwest Asia: La Nina, the warm pool, and Indian Ocean precipitation. *J. Climate* **15**, 697–700.

Berri, G. & Flamenco, E. (1999) Seasonal volume forecast of the Diamante River, Argentina, based on El Niño observations and predictions. *Water Resour. Res*. **35**, 3803–3810.

Hoerling, M. & Kumar, A. (2003) The perfect ocean for drought. *Science* **299**, 691–69.

Kalnay, E. *et al.* (1996) The NCEP/NCAR 40-year reanalysis project. *Bull. Am. Met. Soc.* **77**, 437–471.

New, M., Hulme, M. & Jones, P. (2000) Representing 20th century space-time climate variability. II Development of 1901–1996 monthly terrestrial climate fields. *J. Climate* **13**, 2217–2238.

Schär, C., Vasilina, L., Pertziger, F. & Dirren, S. (2004) Seasonal runoff forecasting using precipitation from meteorological data assimilation systems. *Hydromet.* (in press).

Tippett, M. K., Barlow, M. & Lyon, B. (2003) Statistical correction of Central Southwest Asia winter precipitation simulations. *Int. J. Climatol.* **23**, 1421–1433.

Tippett, M. K., Goddard, L. & Barnston, A. G. (2004) Statistical-dynamical seasonal forecasts of central southwest Asia winter precipitation. *J. Climate* (submitted).