Seasonal forecasting of reservoir inflows in data sparse regions

PLEASE CITE THE PUBLISHED VERSION

PUBLISHER

© Sam Dixon

PUBLISHER STATEMENT

LICENCE

REPOSITORY RECORD

Seasonal forecasting of reservoir inflows in data sparse regions

by

Sam Dixon

A Doctoral thesis submitted in partial fulfilment of

the requirements for the award of

Doctor of Philosophy of Loughborough University

02nd November 2017

Table of Contents

List of acronyms

CA	Central Asia
ENSO	El Niño Southern Oscillation
NAO	North Atlantic Oscillation
DMI	Dipole Mode Index
ISM	Indian Summer Monsoon
GPCC	Global Precipitation Climatology Centre
KNMI	Koninklijk Nederlands Meteorologisch Instituut
GISTEMP	GISS Surface Temperature Analysis
NOAA	National Oceanic and Atmospheric Administration
APHRODITE	Asian Precipitation- Highly-Resolved Observational Data Integration Towards Evaluation of water resources
CMAP	CPC Merged Analysis of Precipitation
TRMM	Tropical Rainfall Measurement Mission
GPM	Global Precipitation Measurement
NCEP	National Centers for Environmental Prediction
NCAR	National Center for Atmospheric Research
NASA	National Aeronautics and Space Administration
IMERG	Integrated Multi-Satellite Retrievals for GPM
TMPA	TRMM Multi Satellite Precipitation Analysis
TRMM RT	TRMM Real Time
JAXA	Japan Aerospace eXploration Agency
CAMS	Climate Assessment and Monitoring System
DMSP	Defence Meteorological Satellite Program
SST	Sea Surface Temperature
ECMWF	European Centre for Medium-range Weather forecasts
GloFAS	Global Flood Awareness System
ESP	Ensemble Streamflow Prediction
AISHF	Automated Information System of Hydrological Forecasts
WBM	Water Balance Model
SRM	Snowmelt Runoff Model
GRDC	Global Runoff Data Centre

TOVAS	TRMM Online Visualisation Service
IOD	Indian Ocean Dipole
PC	Principal Component

List of Figures

x

List of Tables

Abstract

Management of large, transboundary river systems can be politically and strategically problematic. Accurate flow forecasting based on public domain data offers the potential for improved resource allocation and infrastructure management. This study investigates the scope for reservoir inflow forecasting in data sparse regions using public domain information. Four strategically important headwater reservoirs in Central Asia are used to pilot forecasting methodologies (Toktogul, Andijan and Kayrakkum in Kyrgyzstan and Nurek in Tajikistan). Two approaches are developed. First, statistical forecasting of monthly inflow is undertaken using relationships with satellite precipitation estimates as well as reanalysis precipitation and temperature products. Second, mean summer inflows to reservoirs are conditioned on the tercile of preceding winter large scale climate modes (El Niño Southern Oscillation, North Atlantic Oscillation, or Indian Ocean Dipole). The transferability of both approaches is evaluated through implementation to a basin in Morocco. A methodology for operationalising seasonal forecasts of inflows to Nurek reservoir in Tajikistan is also presented.

The statistical models outperformed the long-term average mean monthly inflows into Toktogul and Andijan reservoirs at lead times of 1-4 months using operationally available predictors. Stratifying models to forecast monthly inflows for only summer months (April-September) improved skill over long term average mean monthly inflows. Individual months Niño 3.4 during October-January were significantly ($p < 0.01$) correlated to following mean summer inflows Toktogul, Andijan and Nurek reservoirs during the period 1941-1980. Significant differences ($p < 0.01$) occurred in summer inflows into all reservoirs following opposing phases of winter Niño 3.4 during the period 1941-1980. Over the period 1941-2016 (1993-1999 missing), there exists only a 22% chance of positive summer inflow anomalies into Nurek reservoir following November-December La Niña conditions. Cross validated model skill assessed using the Heidke Hit Proportion outperforms chance, with a hit rate of 51-59% depending upon the period of record used. This climate mode forecasting approach could be extended to natural hazards (e.g. avalanches and mudflows) or to facilitate regional electricity hedging (between neighbouring countries experiencing reduced/increased demand). Further research is needed to evaluate the potential for forecasting winter energy demand, potentially reducing the impact of winter energy crises across the region.

Acknowledgements

Foremost, I would like to express my gratitude to my supervisor Professor Rob Wilby, for his support, patience and guidance throughout the PhD. His knowledge, insight and keen eye for detail has taught me much, and without which the thesis would not be what it is. His prioritisation of me over countless other pulls on his time, in particular during the final stages of writing up, was of great help. Thank you, Rob.

My sincere thanks also go to my progression assessors Professors Helen Rendell and Paul Wood for their insightful comments.

Thanks also go to my friends who have provided welcome distraction throughout the three years. Finally, my thanks go to Jo and my family, for being there throughout.

1 Introduction

Over two thirds of the Tajik population suffer from electricity shortages each winter (World Bank, 2013a). Supply shortages result in electricity rationing every year in Tajikistan (Radio Free Europe, 2016) and increasingly in Kyrgyzstan too (World Bank, 2015). With 70% of Tajikistan's electricity generated by Nurek reservoir (Asian Infrastructure Investment Bank, 2017) and more than 90% of Kyrgyzstan's electricity from the Naryn cascade (including Toktogul reservoir) (World Bank, 2008), winter energy crises are directly linked to reservoir management. Consequently, reservoir operations are inextricably linked to economic development (World Bank, 2013b), regional stability (International Crisis Group, 2014) and public health (World Bank, 2013b; Kalybekova, 2014). The energy crisis was particularly acute during winter 2007/8 when water levels in reservoirs were critically low, resulting in extreme rationing of electricity. Less than 20% of Tajikistan's hospitals and health clinics had access to electricity, resulting in at least 200 infants dying (Unicef, 2008; Laldjebaev, 2010). The real impact was likely much worse, but figures are difficult to obtain (Antelava, 2008). Climate change is expected to exacerbate water stresses in the region, but improved management of water and energy resources has the potential to offset such pressures (Siegfried et al., 2012; Sorg et al., 2012; Sorg et al., 2013).

1.1 Purpose of seasonal forecasting

The long-term safety and efficient operation of reservoirs are vulnerable to hydrological uncertainty. Forecasting expected inflows can reduce vulnerabilities by facilitating improved decision making through increased knowledge about future water availability. For example, drawdown of reservoir levels can be undertaken if high inflow volumes are expected. Long-lead time river flow forecasts can facilitate more efficient reservoir management, with potential annual economic increases of $153 million demonstrated for the Columbia River in the Pacific Northwest of America (Hamlet et al., 2002; Maurer and Lettenmaier, 2003). An integrated management strategy involving multiple interests (e.g. hydropower generation and irrigation) for trans-national river basins could help to reduce tensions surrounding the water-energy nexus of Central Asia (CA). Such a strategy would require improved monitoring systems and integrated flow forecasting

(Clausen et al., 2014). It should be noted, however, that a skilful seasonal forecast does not necessarily improve reservoir management (Yao and Georgakakos, 2001). For improvements to be realised, reservoir managers must be able to interpret and implement forecast information effectively. Insufficient human and institutional capabilities can hinder forecasts being integrated into decision making protocols, a known issue for probabilistic forecasting techniques (Crochemore et al., 2016).

1.2   Challenges of seasonal forecasting in data sparse regions

Seasonal forecasting of reservoir inflows begins with knowledge of regional hydro-climate controls (EMCWF, 2017). This must be acquired via monitoring of the environment before forecasting can commence. The amount of information available varies dramatically between regions. At one extreme, some countries have access to real time information about precipitation, river flow and snow cover (e.g. Norway[1]); whilst others have to rely on manually recorded observations from antiquated hardware (e.g. Kyrgyzstan; World Bank, 2012). Precipitation is the most important atmospheric input for hydrological models, therefore accurate precipitation measurements are vital for reliable flow prediction (Su et al., 2008). In many areas (including the mountainous regions of CA), accurate, real-time, ground-based precipitation observations are scant or non-existent (Artan et al., 2007). Trans-national river basins have additional issues surrounding data sharing, further hindering operational flow forecasting (Hossain et al, 2007). However, the viability of seasonal forecasting is ultimately dependent upon potential predictability, for example from large scale climate modes, which vary in time and space. These limitations, as well as a lack of technical capital in some areas, mean that operational seasonal flow forecasting is a non-trivial task in data sparse regions.

1.3   Opportunities to develop seasonal forecasting

Traditional river flow forecasting techniques rely on accurate ground based measurement of precipitation from gauging stations – a requirement that is often difficult to meet (Artan et al., 2007). However, recent advances in precipitation estimation (from satellites and re-analysis products for example) can provide area-based information, a

---

[1] https://www.nve.no/hydrologi/hydrologiske-data/

distinct advantage over a porous ground-based network. Several products are also available in near real-time, allowing operational flow forecasts to be produced, bypassing the relatively long time it can take to collect and process manually obtained, land-based precipitation data (Collischonn and Pante, 2011). Products can also provide catchment to regional scale coverage of precipitation, combatting issues surrounding data sharing, important for downstream areas within trans-national river basins (Balthrop and Hossain, 2010). Such advantages have attracted attention in recent years from researchers and national agencies alike, with operational flood forecasting models based on such products now in use at both global (e.g. Global Flood Awareness System[2]) and basin scales (e.g. Fan et al., 2016).

Large scale climate modes such as the El Niño Southern Oscillation (ENSO) are known to influence regional precipitation and discharge (Redmond and Koch, 1991; Chiew et al., 1998; Cayan et al. 1999; Dai and Wigley, 2000; Chiew and McMahon, 2002; Emerton et al., 2017). Such relationships, when of sufficient stationarity and practical significance, can be leveraged for seasonal forecasting (Hamlet and Lettenmaier, 1999). Climate mode-streamflow relationships are currently used within operational hydrologic forecasting systems such as BC Hydro (Sene, 2016) and the US Department for Agricultures Natural Resources Conservation Service[3]. Forecasts have been produced via linear and non-linear relationships between climate modes and streamflow (Barlow and Tippet, 2008) and mode-based weighting systems (Werner et al., 2004). Statistical forecasting techniques based on remote climate modes offer significant potential for regions with sparse ground-based monitoring networks, like in CA. Indices of climate modes are available in near real-time, allowing forecasts to be produced before observational data has been collected. Furthermore, where climate modes are related to winter snowpack accumulation, forecast outlooks of summer melt flows can be produced at relatively long lead times (>4 months) (Hamlet and Lettenmaier, 1999).

---

[2] http://globalfloods.jrc.ec.europa.eu/
[3] https://www.wcc.nrcs.usda.gov/wsf/

1.4 Research aim and objectives

The overall aim of this research is to:

- Develop hydrometric data streams and model structures to facilitate reservoir inflow forecasting in data sparse regions using public domain predictors.

The six associated objectives of this research are to:

1) Collate public domain hydro-climatic information for selected river basins and major reservoirs in Central Asia as well as selected climate mode indices;

2) Assess the accuracy of selected information sources via comparison with observed data;

3) Evaluate the effect of varying predictor lead times, locations and averaging periods on forecasts of reservoir inflows in Central Asia;

4) Build and evaluate statistical models using operationally available predictors to forecast reservoir inflows at lead times of 1-4 months;

5) Explore the potential for forecasting reservoir inflows by conditioning mean summer discharge on prior winter mode(s) of climate variation;

6) Evaluate the transferability of statistical and climate mode based models outside the hydro-climate of Central Asia by implementing them in a test catchment in Morocco.

1.5 Thesis structure

Figure 1.1 provides an overview of the steps taken to address the above aim and objectives. The next chapter sets out the context and need for the research. This includes an overview of previous research upon which this thesis builds and discusses available public domain information sources for CA. Chapter 3 introduces the study area and reservoirs used to pilot seasonal forecasting approaches, including information on installed capacities and relative source contributions from ice, snow and liquid precipitation to inflows. The chapter closes with an evaluation of gridded precipitation and temperature products via comparison with observed data. The two seasonal reservoir inflow forecasting methodologies are then described in Chapter 4. First, a statistical forecasting approach is developed using near real-time gridded precipitation and temperature products. Second, a climate mode based approach is devised that

conditions summer reservoir inflows on prior climate mode index phases. The results of both approaches are described in Chapter 5. Chapter 6 then discusses the results, comparing both approaches with previous research. Physical mechanisms are put forward to explain the statistical results, drawing on understanding of the regional hydro-climatology. The transferability of the approaches outside of CA is assessed, giving insight to the drivers of predictability in both CA and North Africa. The scope for operationalising the most promising seasonal forecasting approach is then considered. A review of barriers to implementation is undertaken, before a prototype step-by-step methodology for implementing the forecasts in CA is presented. Finally, Chapter 7 summarises the main contributions of the research to knowledge and proposes several opportunities for further research.

Figure 1.1: **Proposed workflow to address the aim and objectives of the thesis**

2 Literature Review

First, the hydro-political context of Central Asia is provided, stressing the importance of the water-energy nexus on the politics of the region. A summary of the hydro-climate of the region is then provided, giving insight into potential sources of reservoir inflow predictability. Public domain information sources are then identified, before detail is provided on selected products. Relevant research is reviewed regarding seasonal discharge forecasting in the region, placing the study within the context of the current body of literature. A summary of key themes concludes the chapter. The purpose of this Chapter is to situate the region, associated public domain information sources and research objectives within the broader literature.

2.1 A brief history of water management

CA (Figure 2.1) is defined herein as the five 'Stans' of the former Soviet Union: Kazakhstan, Kyrgyzstan, Uzbekistan, Tajikistan and Turkmenistan. Water has been managed in the region for thousands of years. In the interests of brevity, a focus is given to the last century. However, this must be set within the historical context of water management in the region. The Amu Darya once flowed directly into the Caspian Sea, until during the early Holocene the river turned Northwards creating the 'Great Aral Sea'. This was associated with the greater discharges of both the Amu and Syr Darya caused by the warmer and wetter conditions during the Lavlakansky pluvial period (from about 9000 BP) (Kvasov and Mamedov, 1991). The higher lake levels meant water once more discharge from the Aral to the Caspian Seas, until the drier climate around 3500 BP meant this connection ceased. However, even during this period, human influence is thought to have had a major impact on this cessation (Kvasov, 1959). Human influence is not new in the region, with the Horezm population known to have controlled the flow of the Amu Darya from around the 5th century BC (Tolstov, 1962). Overall, the Aral Sea should be viewed as geologically young, and its shallow nature means it has long been susceptible to minor changes in the regions climate as well as human influence (Boomer et al., 2000; Austin et al., 2007).

Following the 1917 Bolshevik revolution, CA became a single administrative unit under Soviet control. In the 1920s, five Republics were created in such a way as to maintain

the requirement of a strong centralising influence from Moscow (Breu and Hurni, 2003). CA received substantial economic and social benefits during the Soviet rule, including subsidies for energy, food and infrastructure. Large scale irrigation was significantly extended throughout lowland CA. Uzbekistan became the cotton basket of Soviet Union and saw irrigated lands increase from ~1.3 million hectares in the 1900s to ~8.5 million hectares currently (Abdullaev and Rakhmatullaev, 2013; Conrad et al., 2016). To this end, several reservoirs were built in the headwaters of the Syr and Amu Darya Rivers to regulate river flows (Wegerich, 2008). Figure 2.2 shows the major dams and all cities located within the study region. A Moscow controlled, unified water-energy system was established. Water from reservoirs was released primarily for irrigation of downstream areas during the summer, and only during peak power demand was water released for hydropower generation. In return, upstream Republics received energy in the form of coal and gas during the winter. This system worked well as both energy and water were managed centrally by Moscow (Sehring and Diebold, 2012). The dissolution of the Soviet Union in 1991 brought independence to the five Republics, as well as an end to Soviet subsidies (Breu and Hurni, 2003).

Figure 2.1 Geopolitical features of Central Asia

Figure 2.2: Location of Dams and cities within the study region

Since independence CA has witnessed radical political changes. Kyrgyzstan is currently the only democratic state following a coup in 2010 which saw the displacement of thousands of Uzbeks (Harding, 2010). The economic situation of Kyrgyzstan and Tajikistan has improved in recent years but is characterised by poverty with dependence on external support, especially in the marginal mountainous areas (World Bank, 2013b). Both the political and economic situation has been aggravated by decaying infrastructure, poor governance and a lack of regional cooperation. On top of these challenges, the region also faces a range of geophysical and meteorological hazards. Earthquakes, avalanches and landslides are common in the Tien Shan and Pamir mountains, with average annual fatalities from natural disasters typically more than 400 people in Tajikistan alone (UNDP, 2012). The primary natural hazard in the region is fluvial flooding, stemming from hydrological variability but also poor dam management, such as flooding of the Arnasay depression due to increased winter release from Toktogul (UNDP, 2012).

Controversy surrounds the management of CA's rivers due to the competing water use demands of upstream and downstream nations. Upstream nations wish to release water for hydropower generation during the winter whereas downstream nations require water for irrigation during summer. Post-independence, energy prices rose dramatically causing upstream nations to act increasingly in their own interests, releasing more water for power generation during winter (Sovaskul et al., 2003). It should be noted that scarcity of water is generally not an issue; rather it is the geographic, seasonal and inter-annual variability of river flow that is problematic (Kraak, 2012). Discussions to improve the situation are set within wider political policies regarding the nation building following the collapse of the Soviet Union (Cruz-Del Rosario, 2009). Current plans to develop new headwater hydropower reservoirs (Kambarata on the Syr Darya, Rogun on the Amu Darya) are closely linked with wider issues regarding legitimacy of political elites and national identity (Menga, 2015). A as well as strengthening regional cooperation, upgrading hydrometeorological networks is crucial for progress, requiring significant investment (UNDP, 2012). Efforts are currently being made to enhance the hydro-meteorological networks in Kyrgyzstan and Tajikistan, but it will take time to build long term accurate records (Schöne et al., 2013; World Bank, 2017). Freely available, public domain information, such as satellite precipitation estimates, could provide information in the near term and be used alongside observational data to provide a more detailed picture of CA hydro-climate.

2.2   Regional hydro-climatology

From the perspective of water resource prediction and management, an understanding of the influence of regional atmospheric circulation patterns on CA climate and the key processes underlying this influence are crucial for accurate seasonal forecasting of river flow. Aizen et al. (1995) identified four key atmospheric circulation patterns that affect CA climate. These were later refined by Inagamova et al. (2002) and Schiemann et al. (2008), and are summarised in Figure 2.3.

Figure 2.3 Schematic of key atmospheric circulation patterns affecting the climate of Central Asia (adapted from Schiemann et al., 2008)

The mountains of the Hindu-Kush, Pamirs and the Karakoram form a barrier to the intrusion of air masses from the South (Schiemann et al., 2007). How effectively this de-couples the climate of CA from the sub-continental climate and the associated monsoon summer precipitation has been of interest over recent decades (Aizen et al., 1995; Webster et al., 1998; Inagamova et al., 2002; Aizen et al., 2009; Pohl et al., 2015). Schiemann et al. (2007) found no spill-over of precipitation into CA during strong Indian Summer Monsoons (ISM). A positive temperature signal was, however, found during strong ISM, particularly across the Amu Darya basin. Schiemann et al. (2007) hypothesised that increased runoff in the Amu Darya found during ISM was caused by the temperature anomaly rather than increased precipitation. However, Arushev et al. (1977) concluded that air masses from the ISM penetrate through the Broghol pass into the southern Pamirs, resulting in increased cloudiness and snowfall during summer causing a lowering of the firn line. Other literature suggests that increased precipitation over the southern Pamirs occurs in June and July during strong ISM, but no evidence is provided to support this view (Aizen et al., 2009; Pohl et al., 2015). Further research is

required to confirm the influence of ISM on precipitation within CA, and particularly how far north its effects penetrate. Furthermore, the link between ISM and summer runoff in the Amu Darya requires further exploration to determine the level of stationarity and underlying physical mechanisms (temperature/melt or precipitation driven).

In general, the precipitation regime of CA reflects the boundary between temperate and subtropical climate zones (Aizen et al., 2001). The spatial variation of precipitation within CA is determined by extreme continentality and topography (Schiemann et al., 2008). According to the Global Precipitation Climatology Centre (GPCC) gridded gauge-based precipitation archive, the lowest precipitation totals (<20mm per month) occur over the Turanian plains in the north-west of CA as well as the Taklamakan desert in the east (Figure 2.4). Generally, a significant increase in precipitation is observed on western and south-western slopes that are exposed to the inflow of moist air masses from the west (>50mm per month according to GPCC) (Figure 2.4). An example of which are the western slopes of the Pamir Mountains located in western Tajikistan. Conversely, interior lee areas such as the Central Tien Shan and the eastern Pamirs receive much less precipitation (<20mm per month, Figure 2.4). The influence of topography and elevation is paramount in understanding the precipitation distribution at local scales (Schiemann et al., 2008). For example, the topography of the Fergana valley is such that westerly air masses converge at its entrance, enhancing precipitation there but resulting in a pronounced minimum within the interior (Schiemann et al., 2008).

Figure 2.4: Elevation model of Central Asia (top plate) and GPCC average monthly precipitation (mm) for years 1981-2010 (bottom plate). Boxes correspond to the same areas as Figure 2.5.

A strong annual regime is observed in the precipitation of CA. Table 2.1 shows that precipitation in CA can be understood by considering the evolution of the jet streams' location throughout the year (Aizen et al., 1995; Chanysheva et al., 1995; Bothe et al., 2012). The annual regime varies spatially within CA, and is more pronounced over the mountainous areas compared with the lowlands (Aizen et al., 1996; Schiemann et al., 2008). Figure 2.5 depicts the variation in annual precipitation regime across CA for six selected regions, an important factor that should be considered alongside inter-annual variability. Relatively little information is available on the spatial-temporal

Table 2.1 Typical precipitation regime of Central Asia (adapted from Schimann et al., 2008)

Months	Precipitation observed
June, July, August	Strong heating of the land surface causes a thermal low to form, resulting in clear, dry conditions over both mountainous and the plains. Periodic intrusions of cooler air masses from the north, north west or west result in continued dry conditions over the plains but precipitation via thunderstorms over the mountains. The jet stream is situated north of CA.
September	Thermal lows rarely form. Dry and clear conditions generally remain due to the southern periphery of the Siberian high being located over CA.
October, November, December	The jet stream moves to the south of CA during October. The cyclonic circulation is activated resulting in less stable weather conditions. To the south of the jet stream, south westerly cyclones carry moist air masses from the east of the Mediterranean to CA. Cyclonic activity continues until mid-January.
January, February	The jet stream moves further south until it is located over Iran, Afghanistan and Pakistan. Southerly cyclones only reach southern areas of CA. The northern half of CA is strongly under the influence of the Siberian high causing calm and clear weather to dominate.
March	The jet stream moves north to the southern periphery of CA. Wintery conditions remain in Kazakhstan and Siberia. In the South of CA a short subtropical spring begins in Iran and Afghanistan. These factors cause a temperature gradient around the jet stream, strengthening cyclonic activity.
Late March, April	Typically months with the highest amount of precipitation, usually falling in showers and thunderstorms. The jet stream moves north across CA.
May	Cyclonic activity weakens, causing a reduction in precipitation.

regimes of snow/rain in CA, with some studies using simple algorithms to predict whether precipitation will fall as snow or rain (Aizen et al., 1995). However, a reduction in the snowfall to precipitation ratio (i.e. decreasing snowfall relative to precipitation) has

occurred over the western Tien Shan Mountains of China between 1961-2010 (Guo and Li, 2015). Altitudes of 1500-2500m have seen the greatest reduction in snowfall to precipitation ratio, with trends less clear outside of this range.

Figure 2.5: Precipitation regime across Central Asia. Bar charts show average monthly precipitation across each of the six areas using GPCC precipitation for the years 1970-2010. Error bars show interquartile range.

The temperature regime of CA is strongly continental, with hot summers and cold winters (Aizen et al., 1996). Very little literature is available on the variation of temperature regime spatially within CA, yet research has been undertaken on the effect of the Indian Summer Monsoon (ISM) on temperature in South Eastern CA. Schiemann et al. (2007) found that during years of strong ISM the mountainous areas to the north and north-west of the sub-continent displayed increased surface temperatures. This was attributed to the propagation of Rossby wave trains excited by latent heat release during monsoon rains (Schiemann et al., 2007).

The Aral Sea basin is mainly fed by the Syr Darya and Amu Darya Rivers (Figure 2.1). Their headwater regions are located in the high mountains of the Tien Shan and Pamir ranges (Figure 2.3). The hydrographs of both rivers are dominated by the meltwater component, with summer flow dependent upon releases from stores of winter snow and glacial ablation (Schiemann et al., 2007; Savoskul and Smakhtin, 2013). The Amu Darya has a greater contribution of glacier melt compared to the Syr Darya (with 23%

and 9% respectively). Both the Amu (45%) and Syr (70%) Darya are heavily reliant on seasonal snowmelt for their annual flow (Savoskul and Smakhtin, 2013).

CA has been described as an area of unambiguous warming, with most rapid air temperature rises observed since the 1970s (Seigfried et al., 2012; Sorg et al., 2012). Winter temperature increases are reported to be greater than in summer over Tajikistan and Kyrgyzstan (Bolch, 2007; Wilby et al., 2011a; Kreigel et al., 2013; Unger-Shayesteh et al., 2013). Observed changes in precipitation are less coherent than those of temperature. Studies have found no significant trend in precipitation anomalies since the 1950s across CA (Bolch, 2007; Kutazov and Shahgedanova, 2009; Unger-Shayesteh et al., 2013). It is also unclear if any changes in seasonal cycle have occurred. Greater precipitation during winter (Aizen et al., 1997; Zhang et al., 2009) has been found for the Tien Shan during the last 50 years.

Detection of trends is hampered by lack of surface data in the post-Soviet period. Using the KNMI Climate Explorer[4] tool an overview of temperature and precipitation changes for the region was undertaken. The KNMI tool uses the GISTEMP[5] surface temperature estimate, which interpolates meteorological station data primarily from NOAA GHCN v3 dataset onto a global 2° grid. Precipitation data is provided by the ERA-20C re-analysis[6], a gridded dataset of ~125km^2 spatial resolution. It should be noted that the significant reduction in meteorological stations post 1991 will likely impact both products. Figure 2.6 and Figure 2.7 show an increase in annual temperature for all countries from the 1970s and interestingly shows a pause in annual temperature rise from ~2000 onwards.

Summer temperatures in Tajikistan and Uzbekistan have continued to rise during this period, but mean winter temperatures have fallen by over 0.5°C. This trend has not been reported elsewhere, and must be treated with caution due to the large inter-annual variability and uncertainty in GISTEMP (Figure 2.6). Overall, annual precipitation totals are similar in 2010 compared with 1900 (Figure 2.7).

---

[4] https://climexp.knmi.nl/
[5] https://data.giss.nasa.gov/gistemp/
[6] https://www.ecmwf.int/en/research/climate-reanalysis/era-20c

Figure 2.6: Annual, summer and winter half year mean temperature anomalies (GISTEMP) and % precipitation anomalies (ERA-20C re-analysis) relative to 1961-1990 for Tajikistan. The 10 year running mean is shown in red. (Source: KNMI Climate Explorer)

Figure 2.7: As for Figure 2.6 except for a comparison of the 10 year running means of Tajikistan, Kyrgyzstan and Uzbekistan

However, cycles of relatively high and low precipitation have occurred. This is particularly clear during summer, where a ~20 year cycle is present since 1930 of positive followed by negative anomalies. Winter precipitation also shows variation through time, with a noticeable dry period for all countries in the late 1970's.

Several studies have projected increases in temperature across CA of ~2°C by the 2050s and ~4-5°C by 2100 (Wilby et al., 2011a; Sorg et al., 2012; IPCC, 2013; Kure et al., 2013; Lutz et al., 2013). Large uncertainty surrounds future changes in precipitation, including which areas could receive more or less precipitation and the timing of this precipitation, yet any changes that do occur are expected to be modest (IPCC, 2013; Kure et al., 2013, Lutz et al., 2013; Sommer et al., 2013). Increased winter precipitation of ~17% is expected by the 2050s (Wilby et al., 2011a; Kure et al., 2013). Despite the possibility of more winter precipitation across CA, Kure et al. (2013) projected decreasing snowfall due to higher temperatures. These intra-annual variations in precipitation, as well as glacier melt, could have significant effects on future snow accumulation and discharge regimes.

An important and currently debated research question is when the 'tipping point' of flows will occur, when increasing glacier melt will begin to decrease due to glacier shrinkage (Sorg et al., 2014a). In the northern Tien Shan it is expected that peak flows could be achieved either in the 2020s or that flows will gradually decrease from present (Sorg et al., 2014b; Gan et al., 2015). However, Kure et al. (2013) project the tipping point to occur between 2060 and 2080 for the Vakhsh River in the Pamirs. Research consistently shows a seasonal shift in peak flows to earlier in the year (Hagg et al., 2013; Kure et al., 2013; Sorg et al., 2014b), estimated to be between 30-60 days for the Syr Darya by 2050 (Siegfried et al., 2012). This shift is attributed to earlier onset of snow and glacier melt driven by increased temperature (Kure et al., 2013).

2.3   Public domain hydro-climatic information

Sources of public domain hydro-climatic information include land-based observations, re-analysis and modelled products as well as satellite estimates. Each product is associated with strengths and weaknesses, meaning that no single product is best for all applications. Preliminary considerations when choosing a product include spatial and

temporal resolution and domain, latency and ease of procurement. Inventories of precipitation products for CA are provided in Table 2.2 and Table 2.3.

2.3.1 Estimated precipitation products

Rain gauge data have several benefits. Relatively long records (from early the 1900s) are available for many regions, and advances in automatic weather station technology allow data to be retrieved in real-time. However, in low capacity areas such as Tajikistan and Kyrgyzstan gauges are generally monitored manually, meaning delays in data retrieval and an increased potential for human error (Wilby et al., 2017).

Known problems with gauge measured precipitation include gauge under catch, particularly during snowfall and questions surround gauge representativeness due to the tiny measured area (Sevruk et al., 2009). The point source nature of gauged precipitation means several gauges may be required to estimate the precipitation distribution of a catchment. The number of gauges will depend upon catchment topography, precipitation regime and area of the basin. Often, gauges are located close to towns or villages, resulting in an under-representation of information for high-elevation, headwater catchments (Sene, 2016). This conflicts with the importance of mountainous regions as 'water towers' for downstream regions (Immerzeel et al., 2010). Furthermore, many regions of the globe have sparse coverage of gauges (e.g. there are only 8 WMO meteorological stations in Kyrgyzstan[7]). Due to these constraints, gridded gauge derived products have been used for hydrological applications (e.g. Fekete et al., 2003; Yatagai et al., 2012).

Gridded precipitation products have the advantage of being area based in nature, providing a more complete picture of precipitation over a catchment. However, the interpolation and averaging procedures involved in the conversion adds uncertainty to the final product (McMillan et al., 2012). Satellite-based precipitation estimates began in the 1970s. They can be broadly separated into satellite only and satellite/gauge merged products. Often, products will combine information from several satellite based instruments to improve spatial and temporal coverage and quality (Sene, 2016).

---

[7] https://www.wmo.int/cpdb/kyrgyzstan

Table 2.2: Gauge derived precipitation products for Central Asia

Name	Link	Temporal Coverage	Temporal resolution	Spatial resolution	Latency	Notes
Global Precipitation Climatology Centre (GPCC)	http://www.esrl.noaa.gov/psd/data/gridded/data.gpcc.html	2003-present for first guess	Monthly	0.5° / 1° / 2.5°	5 days for first guess	• First guess product available in near real time but uses a limited number of stations (number unknown) • Monitoring product latency 2 months and full data product updated irregularly
Asian Precipitation-Highly-Resolved Observational Data Integration Towards Evaluation of water resources (APHRODITE)	http://www.chikyu.ac.jp/precip/english/	~1951-2007 (multiple versions)	Daily	0.25° / 0.5°	N/A	• Temporal variation in number of precipitation inputs used • Low density of observation data in some areas (including Central Asia)
CPC Unified Gauge-based Analysis of Global Daily Precipitation	https://www.esrl.noaa.gov/psd/data/gridded/data.cpc.globalprecip.html	1979-present	Daily	0.5°	1 day	• Quality varies with station density • Daily end time varies by country effecting accumulation
CRU TS 3.22	https://crudata.uea.ac.uk/cru/data/hrg/cru_ts_3.22/	1901-2013	Monthly	0.5°	N/A	• Computes a range of variables • Uses substantially fewer variables than GPCC
Daily Observational	https://data.noaa.gov/	Varies.	Daily	Gauge	1 day	• Detailed dataset with

Data: Global Summary of the Day	dataset/global-surface-summary-of-the-day-gsod	1937-present For Naryn And Tien Shan				information on precipitation, temperature, visibility and wind speed. • Large proportion of missing values, especially for precipitation • Low confidence in data quality
Former USSR data archive	http://cdiac.ornl.gov/ndps/ndp040.html	~1925-1989	Daily	Gauge	N/A	• Unknown data quality • Good coverage of CA up until mid-1980's
Central Asia water Info	http://www.cawater-info.net/bd/index_e.htm	1980-present	Unknown	Gauge	N/A	• Requires formal request for data • Range of hydro-meteorological data available • Not available as of 25/05/2017

Table 2.3: Satellite-gauge merged precipitation products for Central Asia

Name	Link	Temporal Coverage	Temporal resolution	Spatial resolution	Latency	Notes
NOAA CMORPH	ftp://ftp.cpc.ncep.noaa.gov/precip/CMORPH_V1.0	1998-Present	Daily	0.25°	Irregular updates	- Recently extended to cover full TRMM/GPM period - Little information regarding update frequency, but not thought to be real-time.
PERSIANN-CDR	http://chrs.web.uci.edu/research/satellite_precipitation/activities00.html	1997-Present	3 or 6 hourly	0.25°	3 months	- Multiple data inputs - Consistent with GPCP monthly estimates - Relies heavily on Infrared data
CPC Merged Analysis of Precipitation (CMAP)	http://www.cpc.ncep.noaa.gov/products/global_precip/html/wpage.cmap.html	1979-Present	Monthly	2.5°	1 month	- Long data period - Input data vary in space and time - Generally greater uncertainty with increasing latitude
Tropical Rainfall Measurement Mission (TRMM)	http://trmm.gsfc.nasa.gov/	1997-Present (expected end date 2018)	3 hourly, Daily and Monthly	0.25°, 0.5° and 1°	2 months	- Possible problems since July 2014 due to fuel shortage resulting in a change in orbit height - Ended data collection on 15 April 2015

Global Precipitation Measurement (GPM)	http://www.nasa.gov/mission_pages/GPM/main/#.VFJI_PmsV8E	IMERG data from Dec 2014	3 hourly, Daily and Monthly	0.25°, 0.5° and 1°	2 months	• Extension of Integrated Multi-satellitE Retrievals for GPM (IMERG) back to the TRMM era planned to happen in late 2017 • Real time since 2015 currently available
GPCP version 2.2 Satellite-Gaug	ftp://precip.gsfc.nasa.gov/pub/gpcp-v2.2/doc/V2.2_doc.pdf	1979-2011	Monthly	2.5°	2 months	• Provides estimates of uncertainty (random errors only)
CAMS-OPI	http://www.cpc.ncep.noaa.gov/products/global_precip/html/wpage.cams_opi.html	1979-present	Monthly	2.5°	5 days	• Merges Climate Anomaly Monitoring system (CAMS) with Outgoing longwave radiation (infra-red) Precipitation Index (OPI) • Infra-red techniques suffer from significant regional and time dependent biases (merged products adjust infra-red inputs using microwave data)

Satellites do not measure precipitation directly, but infer values from a variety of sensors including microwave, radar and infra-red (Huffman et al., 2007). The area based nature of products suits hydrological applications, and recent increases in temporal and spatial resolution (30 minutes, 0.1° for Global Precipitation Measurement, GPM) make satellite estimates a valuable source of information in data scarce regions. Satellite only estimates are often available in near real-time (latency of ~3 hours to 1 day), but provide less accurate estimates than satellite-gauge combined products (latency ~2 months) (Vila et al., 2009). Some products are also limited to ~55° N/S of the equator (e.g. TRMM and PERSIANN). Those which do provide data at higher latitudes are known to have increased inaccuracies outside the tropics (Ebert et al., 2007). The short record length also limits their usability, with satellite based microwave and radar information typically only available from ~2000 onwards. Consideration should also be made for satellite lifespan and planned future missions. If an operational forecast system is reliant on a satellite product, failure or interruption of that data stream may by critical to the forecast systems performance. Generally, satellite estimates perform better when precipitation regimes tend towards deep convection, meaning they are more accurate during summer and at lower latitudes (Ebert et al., 2007; Huffman et al., 2010).

Re-analysis precipitation is a combination of model and measurement, using observations to constrain the dynamical model that optimises between complete spatial coverage and accuracy (Betts et al., 2006). This allows global coverage over long time periods with no data gaps. However, areas of poor quality and quantity of observational data may see reduction in accuracy. Furthermore, observational inputs to the re-analysis are not consistent over time. In addition to observations, model assumptions and parameters also lead to uncertainties in the final product (Bosilovich et al., 2008; Lorenz and Kunstmann, 2012).

This PhD uses Tropical Rainfall Measuring Mission (TRMM) precipitation estimates and the NCEP/NCAR re-analysis 1 gridded precipitation. These products were chosen due to: short latency, ease of procurement, and proven skill at discharge forecasting in CA (Dixon and Wilby, 2016; Barlow and Tippett, 2008). It should be noted that these are used as a proof of concept, and other precipitation products may by more suitable in

other regions and/or time periods. A more detailed review of each precipitation product, associated quality and prior use in CA is provided below.

2.3.1.1 *Tropical Rainfall Measuring Mission satellite precipitation*

TRMM is a joint mission between the National Aeronautics and Space Administration (NASA) and the Japan Aerospace Exploration Agency (JAXA), launched in 1997 (NASA, 2011). TRMM Multi Satellite Precipitation Analysis (TMPA) is a merged satellite product (based on multiple satellite inputs), available in both real-time (satellite only) and as a satellite-gauge combined product. TMPA 3B43 (V7) (henceforth referred to as TRMM) is used here as a 'research grade' product (latency ~2 months) and TMPA 3B42 (V7) real-time (henceforth referred to as TRMM RT) as an operationally applicable product (latency~2 days). The TRMM satellite itself ran out of fuel and burned up in June 2015, but the TMPA precipitation products are still produced using the remaining satellites in the constellation. GPM satellite has replaced TRMM, and produces an equivalent TMPA product entitled Integrated Multi satellitE Retrievals for GPM (IMERG). IMERG data is, however, only available post 2014. A long term dataset (2000-present) combining the two products is planned for release in 2018, but TMPA is currently the most up to date long term product produced by NASA/JAXA.

TMPA is constructed in three or four stages, depending on whether gauge-correction occurs (NASA, 2011). First, microwave precipitation estimates are calibrated and combined; second, infrared precipitation estimates are created with the aid of the calibrated microwave precipitation estimates; third, microwave and infrared estimates are combined; and fourth (for gauge adjusted products), rain gauge data are incorporated into the final estimate. Microwave precipitation estimates are collected by multiple satellites (TRMM, DMSP, Aqua and NOAA) which cover the area between 50°N and 50°S. Infrared data are collected by the TRMM satellite and provide high temporal and spatial coverage. Rain gauge data used in TRMM are obtained from the Global Precipitation Climatology Centre (GPCC) and the Climate Assessment and Monitoring System (CAMS) (Huffman et al. 2007, Huffman and Bolvin 2013). Gauge-correction is undertaken for TRMM by first summing original three-hour values by calendar month. Second, monthly precipitation gauge analysis is used to create a large-scale bias adjustment to the satellite estimates. Last, monthly gauge-adjusted satellite

estimates are combined directly with gauge precipitation via inverse error variance weighting to create the final product (Huffman and Bolvin, 2013).

The error associated with TRMM estimates is somewhat unclear, and has been described as one of the most urgent matters to address if further development of the product is to occur (Scheel et al., 2011). One such uncertainty stems from the 3 hour discrete sampling interval of TRMM resulting in the entire 3 hours being filled by the measured value. This approach assumes that the random error associated with the sampling interval will even out somewhat over the observation period, but it may have significant impacts over short time scales (Scheel et al., 2011; Condom et al., 2011). Furthermore, the changing availability of microwave sensors during TRMM's lifetime influences the resulting product, with gaps persisting in spatial coverage even when all four microwave sensors are available due to limited swath size (Huffman et al., 2007).

Assessing the quality of the TRMM is not straightforward due to several hindrances. Porous, unevenly distributed monitoring networks often result in relatively few gauge stations to 'ground truth' satellite data with (Ebert et al., 2007). There is no universally accepted approach to assess quality (Hossain and Huffman, 2008). If multiple gauge values are available for a single TRMM cell, averaging is often used. This results in a reduction in extreme values and can lead to relatively higher TRMM values (Scheel et al., 2011). Furthermore, it is impossible to track exactly which gauge stations are used in the TRMM gauge calibration procedure, meaning full independency of the validation dataset cannot be assured. However, calibration with gauge data in areas of low station density, such as Central Asia, can be considered quite approximate (Scheel et al., 2011).

The relief of mountainous areas is known to effect microwave signals, hampering TRMM estimation (Matzler and Standley, 2000; Scheel et al., 2011). Infrared data has uncertainties due to omission of the distance between cloud top and land surface, so does not account for evaporation losses below the cloud top (Petty, 2001). Furthermore, inadequacies of snowfall detection result in an assigned rate being applied for all areas (0.1mm hr^{-1} for TRMM) (Huffman and Bolvin, 2017). In regions with frequent snowfall, such as high mountain areas, erroneous conclusions may be drawn if TRMM

precipitation estimates are applied directly without calibration (Gabremichael, et al., 2010).

Several studies have now evaluated TRMM products across CA. TRMM underestimated precipitation in mountainous areas and during heavy precipitation events, possibly due to difficulties in detecting shallow, orographic rainfall (Adler et al. 2003). Aspect was found to be an important factor, with south facing slopes having higher accuracy and correlation compared with north facing slopes. However, there is low confidence in this finding because of limited data for south facing slopes (Ji and Chen, 2012). Low correlations have also been found near large lakes (e.g., Issyk Kul and Toktogul), likely due to contamination of the microwave signal by water bodies and mountains within the sensor footprint (Karaseva et al., 2012; Guo et al., 2015). Strongest correlations were found in the high plateaus, including for the Naryn gauge (Karaseva et al., 2012).

Gauge-corrected TRMM data show significantly less bias than the real-time version (128%-29%) and increased correlation coefficients from 0.58-0.74 when compared to APHRODITE precipitation (Guo et al., 2015). TRMM has been shown to perform better during the winter compared to summer months, yet TRMM RT exhibits opposing skill (Guo et al., 2015; Dixon and Wilby, 2016). This could be due to issues of snowfall detection from the satellite sensors which is corrected during gauge adjustment (Gabremichael, et al., 2010; Hu et al., 2016). Correlations between TRMM and gauge observations have been shown to reduce at elevations above 2000m during winter in the Chinese Tien Shan (Zhao et al., 2015). This again suggests issues with snowfall detection by TRMM, the gauges used, or both. The 0.25° and 0.5° versions of TRMM perform comparably, yet monthly estimates are much better than daily at the Naryn meteorological station in Kyrgyzstan (with correlation coefficients improving from 0.25-0.93) (Dixon and Wilby, 2016). TRMM can also capture the typical above average seasonal precipitation anomalies over south-west CA, performing well compared to other satellite products (Rana et al., 2017). TRMM over-estimated precipitation at Naryn during July, august and September (Dixon and Wilby, 2016). Overestimation was attributed to local heavy precipitation events under southerly monsoon airflows passing through a TRMM cell but not being measured by the gauge.

2.3.1.2 *NCEP/NCAR re-analysis 1 precipitation*

Re-analysis systems use a climate model that is constrained by observations incorporated via data assimilation routines (Bosilovich et al., 2008). Whilst outputs are generally guided by observations, dynamical model uncertainties lead to uncertainty in the resultant products. This system allows information to be provided for extended periods of time with complete spatial and temporal coverage – a significant advantage for hydro-climatological studies (Betts et al., 2006).

NCEP/NCAR re-analysis 1 (henceforth referred to as NCEP) is produced in three stages. Data is first collected and prepared (including quality controlled), then assimilated using the model and finally distributed to users (Kalney et al., 1996). NCEP uses a frozen global data assimilation system as of 1995. All available observations are used at any given time, meaning the data are the most accurate they can be but changes in the observing system will impact the re-analysis output. Variables in the re-analysis are classified according to the amount that they are constrained by observations. Temperature is partially determined by both the model and observations (type B) whilst precipitation is completely determined by the model (type C) (although all variables are subject to the constraint of the assimilation of other variables) (Kistler et al., 2001). Both type B and especially C variables should be used with caution. However, comparison with observations and climatologies suggest they can provide useful information (Kalney et al., 1997). Even when model estimates are biased, the inter-annual variability of type C fields tends to correlate well with independent observations (Kistler et al., 2001). The antiquated data assimilation system, as well as the relatively coarse spatial resolution (2.5°) are the main limitations of the system compared with more recent re-analysis (such as ERA-Interim) (Climate Data Guide, 2017). However, the real-time nature (latency ~1 day) of the system presents a distinct advantage over all other available re-analysis products for operational users (Climate Data Guide, 2017).

NCEP captures the large-scale spatial distribution of precipitation across CA at an annual timescale (Barlow and Tippett, 2008; Schiemann et al., 2008; Zhou et al., 2017). It was shown to reproduce the transition from an almost uniform seasonal cycle in the north-west to the strongly pronounced seasonality in the southern, mountainous regions. NCEP is known to under-estimate mean annual temperature (by ~ 1°C)

compared with gridded observations across the Pamir (Zhou et al., 2017). Zhou et al. (2017) also note that over the Pamir Mountains NCEP over-estimates the annual regime of both precipitation and temperature by around 100%. Barlow and Tippett (2008) found regional scale climate variability is well represented by NCEP during the cold season. Furthermore, inter-annual variability of precipitation was better reproduced during winter and spring over the mountainous regions than over the Turanian plains (Schiemann et al., 2008). This is of importance as winter/spring precipitation has potential as a predictor of summer flows in the major rivers of CA (Schär et al., 2004). NCEP compares favourably with ERA-40 reanalysis and the Climate High Resolution Model (CRHM) derived precipitation when compared to observations over the Tien Shan and Pamir mountains (Schiemann et al., 2008). However, small scale (<40km) topographically induced precipitation is not well represented by NCEP compared with CHRM, likely due to NCEP's coarse spatial resolution (Schiemann et al., 2008) as well as the small number of stations typically included in data assimilation systems for CA (Unger-Shayesteh et al., 2013).

2.3.2 Teleconnection indices

Large scale ocean atmosphere circulation patterns, such as El Niño Southern Oscillation (ENSO) and the North Atlantic Oscillation (NAO), are warming/cooling patterns in the oceans which affect atmospheric circulation and weather patterns around the world. The NAO consists of a north-south dipole of pressure anomalies between the central North Atlantic (centred over the Azores) and Iceland in the North. NAO indices describe relative changes in pressure between these two regions. A positive (negative) NAO index represents lower (higher) pressure in the high latitudes and above (below) normal pressure over the central North Atlantic. Positive NAO phases result in stronger westerly winds bringing increased (reduced) winter precipitation to northern (southern) Europe (Hurrell, 1995). NAO varies considerably at both intra- and inter-annual timescales, although several months with either positive or negative NAO is also common.

ENSO refers to fluctuating Sea Surface Temperatures (SSTs) in the equatorial Pacific Ocean. El Niño (La Niña) has warmer (cooler) than average SSTs across the central and eastern equatorial Pacific. The cycle between El Niño / La Niña usually occurs

every 3-7 years. Several indices are available to define ENSO conditions based on SST anomalies. The multiple indices (1, 2, 3, 3.4 and 4) represent SSTs in different regions (from east to west) across the eastern and central equatorial Pacific.

Large scale climate modes have long been known to influence hydrological variables such as regional precipitation and discharge (Redmond and Koch, 1991; Chiew et al., 1998; Cayan et al. 1999; Dai and Wigley, 2000; Chiew and McMahon, 2002; Werner et al., 2004; Emerton et al., 2017). Such relationships, when of practical significance and sufficient stationarity, can be leveraged for seasonal forecasting (Hamlet and Lettenmaier, 1999, Kennedy et al., 2009). Climate mode-streamflow relationships have been used widely for hydrologic forecasting and are currently used within operational forecasting systems, such as BC Hydro (Sene, 2014) and the US Department for Agricultures Natural Resources Conservation Service[8]. Forecasting techniques based on remote climate modes have significant potential for regions with sparse ground-based monitoring networks, as is the case for much of CA. However, recent research has highlighted the complex nature of ENSO-river flow relationships, being spatially variable, impacted by phase strength (Lyon and Barnston, 2005), interactions with other climate drivers (Emerton et al., 2017) and in their nature (linear/parabolic) (Fleming and Dahlke, 2014). It is important to note that, due to these factors, even in areas with strong relationships it is possible that opposite anomalies to those expected from a particular mode of variability can occur.

CA has received relatively little attention regarding possible links between hydrological variables and large scale climate drivers. This lack of research may reflect the complexity of the hydro-climatology of the region and relative remoteness from major modes of climate variability (Mariotti, 2007). However, a growing number of studies have begun to establish links between the hydro-climate of CA and NAO and ENSO. Both significant (Aizen et al., 2001) and insignificant (Syed et al., 2006; Huang et al., 2013; Barlow and Hoell, 2015) concurrent negative correlations have been found between NAO and CA precipitation during winter. Correlation strengths vary spatially, although are generally weaker across Kyrgyzstan ($r = 0.15$) than Tajikistan ($r = -0.30$) for the years 1951-2000 (Syed et al., 2006). A link between positive NAO and drought in

---

[8] https://www.wcc.nrcs.usda.gov/wsf/

south west CA has been suggested by Barlow and Hoell (2015). Negative correlations have also been found between NAO and mean December-March precipitation and streamflow across the Middle East by Cullen and deMenocal (2000) and Cullen et al., (2002).

More information is available regarding ENSO's relationship with CA precipitation, showing positive concurrent winter precipitation anomalies during El Niño (Dai and Wigley, 2000; Syed et al., 2006; Mariotti, 2007, Yin et al., 2014; Hoell et al., 2017; Rana et al., 2017). Relatively strong correlations have been found during autumn ($r = 0.50$) and spring ($r = 0.40$) during the period 1948-2000 (Mariotti, 2007). A link between drought in south west CA and concurrent La Niña has also been reported (Barlow et al., 2002; Hoerling and Kumar, 2003; Hoell et al., 2014). More recent work shows increased likelihood of abnormally high flows during El Niño across CA (Emerton et al., 2017). An increasing influence on ENSO on winter precipitation has been recorded for northern India as well as a decreasing influence of NAO in recent decades (Yadav et al., 2009). On the other hand, positive temperature anomalies in the southern Pamirs coincide with strong ISM (Schiemann et al., 2007). It is possible that a strong ISM dampens the effect of ENSO on precipitation across CA, as shown for western India and Pakistan (Ashok et al., 2004).

2.4 Seasonal flow forecasting

Seasonal flow forecasting is used to provide estimates of future flow conditions typically at monthly to seasonal granularity. The time interval between issuing the forecast and the beginning of the period of the forecast is called the lead time (WMO, 2000). Lead times encompassed within seasonal forecasting are not clearly defined, but are classed here as between 1-12 months (Sene, 2016). The premise of seasonal forecasting is that more informed decisions can be made if knowledge of future climate conditions can be provided (e.g. drawdown of reservoir levels before a flooding event). Forecast skill originates from several sources, including persistence in initial hydrologic conditions, simple rainfall-runoff relationships, long-range meteorological forecasts and lagged relationships with large scale modes of climate variability (such as the NAO).

2.4.1 Seasonal forecasting approaches

Seasonal forecasting models can be split into two main groups: statistical and dynamical. Dynamical models represent the physical features of a system and have at their core a set of mathematical equations describing the interactions of the earth's energy and moisture states (Mantua et al., 2008). The main advantage of dynamical models is that the influence of a model variable on another variable is guided by the laws of physics. Hence, the model is expected to function under conditions outside those used in the calibration procedure. Their use is hindered, however, by their high computational and data input demands, potentially requiring the approximation of parameters such that their realism and therefore performance is compromised (Mantua et al., 2008). However, dynamical forecasts of climate are provided operationally by several centres (e.g. European Centre for Medium-range Weather Forecasts (ECMWF)[9], International Research Institute for Climate and Society[10], Global Flood Awareness System (GloFAS[11])).

Statistical models can be sub divided into analogue and empirical methods. Analogue approaches look for patterns in historical flow records that are similar to those occurring today and use this to forecast. The underlying assumption is that river flows observed previously that are similar to the recent past will provide valuable information on what flows could occur in the near future. The main source of uncertainty is the initial conditions of the system. Empirical methods such as regression-based techniques (e.g. Archer and Fowler, 2008) relate observations of the variable to be predicted (e.g. river discharge) to observations of one or more other variables (e.g. seasonal snowpack or precipitation totals). They are often preferred due to their computational ease relative to dynamical models and in many cases can provide comparable results (Mantua et al., 2008; Schepen and Wang, 2015). However, statistical models require long homogeneous records to calibrate and verify predictor-predictand relationships (Wood and Lettenmaier, 2006). Statistical models may also assume the climate is stationary, not accounting for anthropogenic impacts or global warming (Hamlet and Lettenmaier, 1999). Furthermore, their lack of physical process representation might compromise

---

[9] https://www.ecmwf.int/en/forecasts
[10] http://iri.columbia.edu/our-expertise/climate/forecasts/seasonal-climate-forecasts/
[11] http://globalfloods.jrc.ec.europa.eu/

their ability to forecast events outside the range of the calibration dataset (Mantua et al., 2008).

Ensemble Streamflow Prediction (ESP) is a widely used technique for seasonal river flow forecasting (particularly in the Pacific NW of America) (e.g. Wood and Lettenmaier, 2006). Ensemble forecasts produce a set of equally-likely outlooks for the future that can be based on historical climate (similar to analogue techniques) or long term forecasts of climate (using dynamical techniques such as numerical weather prediction). This meteorological information is then used to force a hydrologic (or land surface) model. This versatile approach can enable a greater understanding of the initial condition and parameter uncertainty surrounding a forecast compared to deterministic methods, with the aim of facilitating more informed decision making (Cloke and Pappenberger, 2009). A major limiting factor currently is the ability of practitioners to make improved decisions from probabilistic ESP forecasts. This is due to a lack of practitioner 'know-how', particularly when interpreting uncertainty bounds provided within probabilistic forecasts as well as differing attitudes to risk, uncertainty and error between academics and end users (Demeritt et al., 2007; Cloke and Pappenberger, 2009). Furthermore, instigating change within organisations from their current, well understood, tried and tested methodology to a new system is extremely difficult (Demeritt et al., 2010). The more complex and large the shift in practice, the more difficult this becomes (Ludwig, 2009).

With the increase in computing power a shift has been made towards continental and global scale seasonal forecasting systems (Alfieri et al., 2013). These have potential to provide forecasts where there are currently none available or in transboundary basins where floods originate upstream. However, such forecasting systems have coarse resolutions and are designed to perform optimally over large areas. Conversely, basin scale seasonal forecasting allows models to be optimised for individual basins. Small scale nuances can be captured such as hydrologically important areas which could be a focus for enhanced ground based monitoring (Dixon and Wilby, 2016). Basin scale forecasts can also be tuned to fit specific decision makers' requirements, providing outputs in the required format at suitable lead times and level of detail/complexity. This is known to be vital in the successful integration of forecasts into decision making protocols (Demeritt et al., 2010). No single model is perfectly suited to all applications.

Therefore, model selection should be made accounting for the aims of the study, the available data quantity and quality whilst always keeping in mind the intended application.

2.4.2 Hydrological model performance assessment

To evaluate the relative skill of each model across varying lead times it is necessary to assess their performance. This provides information on the relative merits of each model and their ability to seasonally forecast discharge accurately. Krause et al. (2005) asserts that performance assessment provides: (1) quantitative indicators of model skill at reproducing watershed behaviour, (2) means of evaluating improvements in a given modelling system and (3) mechanisms for comparison of results obtained from different models. It should be noted that this section does not assess forecast value, but instead performance (i.e. to identify relative strengths and weaknesses of the model). Initial performance assessment should be carried out subjectively via a visual inspection of the modelled and observed time series. This allows the modeller to evaluate the model for systematic (under-/over-prediction) and dynamic (timing, rising limb, low flows) errors (Dawson et al., 2007). To achieve objective assessment, however, mathematical estimates of error using evaluation metrics should be undertaken (Krause et al., 2005).

A metric can be defined as a "system of parameters, or methods of quantitative assessment, for something that is to be measured" (Dawson et al., 2007, p.1035). Individual error metrics used in hydrology may not always provide a comprehensive assessment of model performance for a specific application, and is therefore of limited use (Teegavarapu and Elshorbagy, 2005). Metrics used in isolation can under-/over-estimate model performance. To combat this, use of multiple error metrics suitable to the intended model application is recommended (Krause et al., 2005; Gupta et al., 2008; Fowler et al., 2012). Dawson et al. (2007) grouped error metrics into three categories: (1) statistical parameters of observed and modelled time series attributes (e.g. maximum, minimum, mean); (2) statistical parameters of the residual error between observed and modelled time series and; (3) dimensionless coefficients that contrast model performance with accepted norms or recognised standards (e.g. Nash-Sutcliffe coefficient). Definitions, along with known strengths and weaknesses of frequently used metrics for evaluating hydrological models, are provided within Dawson

et al. (2007). When choosing metrics the nature of the predictand (continuous/categorical/probability distribution) and forecast attributes to be measured (bias/accuracy/skill) should be considered (Wilson, 2004).

When seasonal forecasting we generally wish to know how the model will perform in practice, rather than how it performs using the training dataset. In particular, we wish to avoid overfitting a model to a training dataset, meaning performance is tailored to using only this dataset. Cross validation can be used to address this issue. The simplest form of cross validation withholds a section of data during training to be used for testing of model performance. However, particularly when using a statistical model, a long data series is required for the predictor-predictand relationship to be fully captured. Furthermore, the estimate of model performance could be misleading if an 'unfortunate' period of data is used for testing. Therefore, leaving out a period of data during training to allow for testing is undesired. A k-fold cross validation addresses this issue by splitting the data into *k*-subsets, and performing the withholding method *k*-times (Figure 2.8). This allows the model to be tested on the full period of available data. The more folds used, the greater the computational demand. Generally, the larger the available data set the fewer number of folds are required to achieve an accurate estimate of model performance (Kohavi, 1995).

Data used to train model   Data used to test model
Figure 2.8: Conceptual figure of 9-fold cross validation procedure

2.5   Operational hydrological forecasting in Central Asia

Previous research into discharge simulation for rivers in CA has been undertaken for several purposes including climate change assessment, improved process understanding, seasonal forecasting and natural resources management. Several

studies have assessed future water resources under various climate change scenarios and adaptation strategies (e.g. Siegfried et al., 2011 and Ismaiylov et al., 2007). Due to the differing aims of these studies, scant information is provided regarding hydrological model performance and inputs used are not available in real time.

Information about operational seasonal forecasting practices in CA is seldom published. However, it is known that some national hydro-meteorological agencies in the region produce forecasts of April-September discharge each month from January until June (Apel et al., 2017). Empirical methods are used to relate precipitation, temperature and snow water equivalent to seasonal discharge, which is sometimes only available in analogue form as look-up tables or graphs. For example, the Automated Information System of Hydrological Forecasts (AISHF) model is used in Uzbekistan (Agaltseva et al., 1997). AISHF is an empirical model using observed meteorological data as well as modelled snowpack and glacier melt to provide operational forecasts using specialist software allowing interactive analysis. The forecast was reported to meet the information needs of the users, however, it is unclear at what lead time satisfactory forecasts were produced as this information was not provided (Agaltseva et al., 1997).

Wilby et al. (2011a) used several models to estimate annual inflow to the Kayrakkum and Nurek reservoirs on the Syr and Amu Darya respectively. They used a Water Balance Model (WBM), the Snowmelt Runoff Model (SRM) and a nonlinear multiple regression model. Model inputs included observed meteorological station precipitation and temperature, as well as remotely sensed snow covered areas. Model performance was found to be better for Kayrakkum than Nurek, likely due to upstream reservoirs smoothing climate related variability (Wilby et al., 2011a). Temperature was successfully used as a proxy for snow and glacier melt in both basins (Wilby et al., 2011b). Following years with higher than average temperature, inflows to Nurek were generally lower. This was attributed to higher evaporative losses and less carry-over of snow and ice between years (Wilby et al., 2011b). This relationship has been reported elsewhere, and has the potential to contain useful information for seasonal forecasting (Schär et al., 2004; Archer and Fowler, 2008). Pereira-Cardenal et al. (2011) utilised TRMM precipitation estimates along with radar altimetry of Toktogul reservoir water levels as a proxy for discharge to attempt to improve real time forecasting. Comparable results were found between modelled inflows and satellite altimetry. However, no

observed data were used to evaluate the model. Recently, synthetic runoff estimates have been derived from gridded discharge by routing land surface/meteorological model outputs (Zaitchik et al., 2010; Khouakhi, 2017). Although in early stages of development, such techniques have potential for discharge estimation in areas lacking hydro-meteorological monitoring.

Some studies have focused on seasonal forecasting of river flows in Central Asia. For instance, Baumgartner et al. (2000) used meteorological forecasts as inputs to SRM. Promising results were obtained for forecasting summer flows (r^2 = 0.74-0.97), but little information is provided regarding model inputs or lead times used to obtain these results. Satellite snow cover data were reportedly utilised by Yakovlev (2005) to forecast flows of the Panj River, Tajikistan. Although availability of cloud free imagery is critical for accurate flow forecasts, little information is provided regarding model performance (referenced within Gafurov, 2010).

Schär et al. (2004) used December-April re-analysis precipitation to forecast May-September withdrawal adjusted flows in both the Syr and Amu Darya rivers. Re-analysis precipitation performed well compared with rain gauge values, due to the more complete spatial coverage by the re-analysis compared with the sparse network of land-based stations. Forecasts performed better for the Syr Darya (r = 0.92) compared to the Amu Darya. Poor performance may have been caused by lower quality re-analysis precipitation over the Amu Darya basin and/or lower quality withdrawal corrected runoff figures. Successful forecasts of summer inflows to the Mangla Dam, Pakistan, from previous winter observed precipitation and temperature were obtained using a parsimonious liner regression model (Archer and Fowler, 2008). Although outside of CA, a similar hydrological regime (dominated by summer snow and glacier melt) is observed. Winter precipitation was found to be a useful predictor of following spring and summer inflows, giving promise to similar methodologies in CA.

Dixon and Wilby (2016) developed a parsimonious multiple linear regression model to forecast inflows to Toktogul Reservoir on the Syr Darya. Inputs included TRMM as well as observed precipitation and temperature and antecedent flows. Hindcast skill was superior to the mean monthly flow for lead times up to three months. Over 80% of the variance in monthly inflows is explained with three-month lead, and up to 65% for

summer half-year average. The findings of Schär et al. (2004) and Dixon and Wilby (2016) give promise for seasonal forecasting but neither study used solely operational precipitation data, meaning that forecasts could not be made in real-time. This is of critical importance if forecasts are to be used to support reservoir operation.

Apel et al. (2017) used operationally available data (previous winter observed precipitation, temperature and discharge as well as remotely sensed snow cover) to forecast April-September discharge in 13 basins in CA for the period 2000-2015 (note that data used is not publicly available). Using automatic predictor selection and multiple linear regression R^2 values of >0.7 were achieved at a lead time of two months (February), rising to 0.86-0.96 for zero lead forecasts (issued April 1st). However, results were based on a very limited number of observations (just 16) with no adjustment for sample size. Furthermore, the statistical significance of results was not provided. The authors attributed their high skill scores to the separation of precipitation (winter maximum) and runoff (spring/summer maximum) regimes.

Further potentially operational seasonal forecasting systems have been developed for the region. For example, Tippett et al. (2005) forecast December-March precipitation from October Pacific SSTs. Significant lagged correlations between SST and precipitation were found across Tajikistan ($r = 0.40$) and western Kyrgyzstan ($r = 0.35$) during the period 1999-2003. Gerlitz et al. (2016) used NCEP re-analysis alongside SST to forecast winter and spring precipitation in CA. Correlations of $r \sim 0.5$ were found between forecast and observed winter-spring precipitation across northern CA at lead time 1.5 months. Correlations were generally stronger for northern CA compared with southern CA. Wet and dry years were well represented by the model, but the variability of precipitation rates was under-estimated. When lead time increases to 4.5 months correlations weaken to $r \sim 0.2$ in northern CA and less in southern CA. Barlow and Tippett (2008) forecast summer (April-August) river flows from winter Pacific SSTs and winter NCEP re-analysis precipitation for the years 1950-1985. They report modest cross validated correlation ($r \sim 0.4$) skill between observed and forecast flows in the Vakhsh at Garm. Typically, stations in mountainous regions had the weakest cross validated correlations ($r < 0.4$). Flow volume was found to be unrelated to model skill. These studies confirm that summer inflows (refill period) to headwater reservoirs can be forecast from winter precipitation/snow accumulation, albeit with varying degrees of

accuracy. Furthermore, relationships between climate modes and winter precipitation in CA gives hope for extended lead time forecasting of summer inflows.

2.6 Summary

This chapter has provided a brief overview of the hydro-politics and hydro-climatology of CA. Key sources of public domain hydroclimatic information and previous flow forecasting studies have been evaluated. It is clear that an improved operational seasonal reservoir inflow forecasting system could benefit the region, with hydro-power and water supply companies as well as farmers being foremost beneficiaries. Due to limited surface observations and significant obstacles to data procurement, near real-time precipitation estimates from satellite and re-analysis sources offer appealing substitutes for ground-based precipitation measurements. Significant concurrent positive correlations have been reported for ENSO and winter precipitation that are particularly strong in autumn and spring. Furthermore, significant lagged correlations have been found between October Pacific SSTs and December-March precipitation, suggesting longer range predictability of winter precipitation. Previous research has shown the possibility of forecasting summer inflows in CA from previous winter precipitation. This suggests that there may be potential for forecasting summer inflows to significant reservoirs using previous winter precipitation and SSTs.

Although dynamical forecasting offers some advantages over statistical techniques, the former have significant data requirements, computer power and are costly to implement. Furthermore, at the time periods and lead times of interest recent research has demonstrated that statistical techniques can provide operationally useful skill in the region. This study will apply statistical techniques to forecast inflows to several headwater reservoirs using operationally available predictors. Information on specific study sites as well as the data products used in the research is provided in the following chapter.

3 Study areas and data

Before seasonal forecasting can commence, an understanding of the target basin(s) context should be established. Appropriate data must then be accrued to facilitate accurate seasonal forecasting. This chapter introduces the reservoirs and their respective basins and identifies the available information to support operational seasonal forecasting of their inflows. Detail is provided regarding reservoir locations and the hydro-climatological context of each site in section 3.2. Section 3 summarises the ground based observational data used. Selection of remotely sensed information is reviewed in section 3.3 before identification and selection of modes of climate variability and their respective indices is given in section 3.4. An evaluation of remotely sensed (TRMM and TRMM RT estimated precipitation) and re-analysis (NCEP precipitation and temperature) products via comparison with observed data is then provided before a review of data selected concludes the chapter.

3.1 Meta information

3.1.1 Reservoir sites

Four reservoirs and their respective headwater basins are used in this study: Toktogul, Andijan Kayrakkum and Nurek (Figure 3.1). The first three reservoirs are located on tributaries of, or on the main stem of, the Syr Darya River. Toktogul and Andijan are located upstream of the inflow to Kayrakkum. Nurek is positioned on the Vakhsh River, a tributary of the Amu Darya. Both the Syr and Amu Darya rivers ultimately flow into the Aral Sea. Toktogul, Andijan and Nurek dams were chosen due to their headwater location and relatively large size compared to other reservoirs in the region. The operation of these reservoirs is therefore critical in determining power generation and flow downstream. Kayrakkum reservoir was selected due to its size and importance in providing irrigation to downstream areas. Differences in average elevation, contribution of snow and/or glacier melt and climatic conditions will also provide insight into the performance of the seasonal forecasting models under various conditions. A summary of each basin's characteristics is provided in Table 3.1.

Figure 3.1: Location of study sites, rivers and meteorological stations

Table 3.1: Summary of river basin charectaristics

	Toktogul	Andijan	Kayrakkum	Nurek
Year of commissioning	1974	1974-84	1956-59	1971-79
Installed capacity (MW)	1200	190	126	3015
Annual mean power generation (MWh)	4,400	172	690	11,400
Meteorological station	Naryn	Uzgen	Uzgen	Sari-Tash
Meteorological station elevation (m)	2039	1014	1014	3155
Average annual precipitation (mm)	315	732	732	420
Basin area above inflow gauge (km^2)	47,052	12,348	124,276	31,059
Average annual inflow (m^3 s^{-1})	445	133	656	662
Wettest month	May	April	April	May
Peak flow month	June	June	December	July
% of annual precipitation occurring between April and September	68	41	41	63
% of annual flow occurring between April and September	77	77	36	83
Annual inflow as a % of annual met station precipitation (TRMM basin average precipitation)	36 (24)	47 (65)	24 (23)	160 (142)

Toktogul

Toktogul reservoir (Figure 3.2 and Figure 3.3) is located on the Naryn River in the central Tien Shan mountain range of Kyrgyzstan. The Naryn is fed by a major tributary below Song Kol Lake which runs during the melt season (April-September) (Dixon and Wilby, 2016). Construction of Toktogul began in 1960 and was commissioned in 1974. The reservoir is operated for both hydropower generation and irrigation. The power station has an installed capacity of 1200 MW and a mean annual production of 4400 MWh (CAWaterInfo, 2017a).

Land cover is mainly grass and shrub, with pastoral farming on the mountain sides. The main meteorological station of the basin is located at Naryn, although a higher elevation station (Tien Shan) is present yet only limited precipitation data could be procured. Naryn is the largest town in the basin, with a population of around 35,000 people. Kambarata II dam lies upstream of Toktogul. Construction of Kambarata II began in the 1980s, but was abandoned shortly after and only recommenced in the late 2000s (Hydro World, 2010). The dam was created by an artificial mass movement in December 2009, before being commissioned in 2012, and has a height of 60m (Havenith et al., 2015). The subsequent filling of the dam is likely to have impacted inflows to Toktogul post 2009, however the extent of effect is unknown. Further dams are planned upstream, but currently have not secured the required funding (Michel, 2016). If built, the Kambarata I dam would surpass Toktogul as the largest installed capacity for hydropower production in Kyrgyzstan (World Bank, 2010).

Andijan

Andijan reservoir (Figure 3.4 and Figure 3.5) is located on the Kara Darya in an area known as the Fergana Range, within the central Tien Shan Mountains. The dam lies within Uzbekistan, although the reservoir extends into neighbouring Kyrgyzstan. Construction began in 1969 and was commissioned in 1974. A second plant below the original was commissioned in 2010. The reservoir is operated for both hydropower generation and irrigation. The power station has an installed capacity of 190 MW and a mean annual production of 172 MWh (CAWaterInfo, 2017a). The headwaters are lower in elevation compared to those of the Toktogul basin reaching a maximum of 3,800m above sea level. The city of Andijan is the largest population centre in the region with

over 300,000 residents, but is located below the dam. Several towns are located around the reservoir, the largest being Uzgen at around 50,000 people. No known impoundments are located above the reservoir.

Figure 3.2: Toktogul dam, Kyrgyzstan[12]

Figure 3.3: Satellite image of Toktogul reservoir (Google Earth)

Figure 3.4: Andijan Dam, Uzbekistan[13]

---

[13] http://test4.ocean.washington.edu/file/Andizhan+Dam

Figure 3.5: Satellite image of Andijan reservoir (Google Earth)

Kayrakkum

Kayrakkum reservoir (Figure 3.6 and Figure 3.7) is located on the Syr Darya River, and receives inflow from the above two basins. The reservoir is located within Tajikistan, but the upstream area includes Kyrgyzstan and Uzbekistan. Construction began in 1952 and was commissioned in 1956. The reservoir is operated for hydropower generation, irrigation, recreation, fish breeding and water supply. The power station has an installed capacity of 126 MW and a mean annual production of 690 MWh (CAWaterInfo, 2017b). The reservoir is located at the lower end of the Fergana valley, one of the most productive regions for cotton growing in Uzbekistan and nearly solely irrigation fed. This has resulted in the region being one of the most populous in Central Asia having around 20% of the region's population in less than 5% of the land area. Agriculture is the dominant land use in this region, much irrigated from the Syr Darya, with crops including cotton and grain. The reservoir inflow regime is dominated by the regulated discharges from the Naryn cascade (consisting of five hydropower plants below Toktogul reservoir) and Andijan as well as withdrawals for irrigation. It is expected that the highly regulated flows, evidenced by the opposing inflow regime compared with the other reservoirs, will not be predicable using the methodology used herein (Figure 3.8). However, the reservoir was included in the study to assess the ability of the model to forecast heavily

regulated reservoir inflows. Figure 3.1 shows the large area below Toktogul and Andijan reservoirs which contributes to runoff. It is plausible that some predictability can be derived from this area as well as outflows from Toktogul and Andijan reservoirs.

Figure 3.6: Kayrakkum hydropower plant (Michael Friedhoff)

Figure 3.7: Satellite image of Kayrakkum reservoir (Google Earth)

Figure 3.8: Average monthly inflow (left) and monthly inflow as a % of annual discharge (right) into the four study reservoirs for years 2001-2010

Nurek

Nurek reservoir (Figure 3.9 and Figure 3.10) is located on the Vakhsh River in the Pamir Mountains of Tajikistan. The headwaters of the Vakhsh rise to an altitude of more than 7,400m. Glacier melt provides a greater percentage of annual discharge compared with the other study basins, reflected in the annual inflow regimes (Figure 3.8) (Savoskul and

47

Smakhtin, 2013). An estimated 26% of annual runoff derives from glacier melt (Kemmerikh et al., 1972). Construction began in 1961 and the plant was commissioned in 1971. The reservoir is operated for hydropower generation and irrigation. The power station has an installed capacity of 3015 MW and a mean annual production of 11400 MWh (CAWaterInfo, 2017b). Nurek supplies over 70% of Tajikistan's total power demand (Hydro World, 2017). No large towns are located within the basin, with Sary Tash being an administrative centre but having less than 2000 residents. Upstream of Nurek is the site of the Rogun Dam which was partially constructed during the 1970s and re-started at the end of 2016 (Hydroworld, 2016). It is unknown whether Rogun's construction has affected the inflow regime of Nurek, but filling of Rogun could begin as early as 2018.

There is generally little information about the source contributions to mean discharge into each of the reservoirs (Unger-Shayesteh et al., 2013). Kemmerikh (1972) estimated the relative contributions of seasonal snowmelt, groundwater, glacier melt and rain to both the Naryn and Vakhsh Rivers (Table 3.2). Note that Kemmerikh's (1972) definition of the Naryn basin differs from this study, only comprising the headwater region.

Figure 3.9: Nurek dam, Tajikistan (Robert Wilby)

Figure 3.10: Satellite image of Nurek reservoir (Google Earth)

Therefore, it would be expected that a greater proportion of glacial melt would be found for this stretch of river compared to the basin definition used here. Savoskul and

Smakhtin (2013) estimated relative contribution to discharge of glacier melt, snow melt, groundwater and rain for the Syr and Amu Darya rivers. They report greater glacial melt contributions for the Amu Darya compared to the Syr Darya (25% and 9% respectively) but the opposite for snow melt (45% and 70% respectively) during the period 1960-1990. These findings as well as the annual inflow regimes shown in Figure 3.8 and elevation of each basin suggest a greater proportion of glacier derived discharge into Nurek than Toktogul, and Toktogul compared with Andijan.

Table 3.2: Estimated contributions to mean annual discharge by snow melt, groundwater, glacier melt and rainfall sources. Adapted from Kemmerikh (1972)

Basin	Basin area (km^2)	Snow melt (%)	Groundwater (%)	Glacier melt (%)	Rain (%)
Naryn*	10,500	32	31	32	5
Vakhsh**	31,200	37	37	26	-

3.2   Ground based information

Using statistical methods for seasonal discharge forecasting requires high quality observational time series that are homogeneous. However, collection of such observational data in CA is hindered by several issues (Unger-Shayesteh et al., 2013). Hydro-climatological data sets at daily or sub-daily resolution are not readily available to international researchers and most long term records are located in the foothills or plains of CA. Furthermore, the network of observational stations was severely degraded after the break down of the Soviet Union in 1991 and several stations have known Inhomogeneities due to station relocation, urbanisation and changes in measurement techniques (Unger-Shayesteh et al., 2013). These factors cause procurement of hydro-climatic data to be challenging, particularly for recent decades. This has resulted in the meteorological data sourced for this study having varying timespans and temporal resolutions.

---

* Note that Kemmerikh's (1972) definition of the Naryn basin equates to the eastern 20% (furthest upstream) of the area used in this study.
** The author believes that rain does contribute towards discharge into Nurek. This table was incorporated as it is one of the few studies to provide contributions for the headwater of the Vakhsh.

There are several sources of bias likely to be present in hydro-meteorological data used in this study (Wilby et al., 2017). For example, site selection can introduce bias by being located close to population centres at generally low elevations, bringing into question the representativeness of a station and the potential for an urban heat island effect (Jones et al., 2008). The condition and calibration of discharge measurement stations is also a concern for the data used here. Discharge is likely measured via stage at locations of highly mobile channel cross sections, meaning the stage-discharge relationship used will quickly become outdated (Unger-Shayesteh et al., 2013). Precipitation data obtained from rain gauges are known to have issues related to gauge-under catch caused by wind, a particular problem during the winter.

Both meteorological and hydrological stations in CA have predominantly been operated manually, requiring an observer to record readings. Data collected in this way are subject to observer measurement biases, for example tending towards labelled/large tick marks on measurement devices or observations not being made during weekends/holidays (Daly et al., 2007). Bias could be introduced during post-processing of data, especially when not obtained first hand. Meta data is useful here, for example in helping to understand if filling of gaps has been undertaken and how this was done. This is of particular importance in a region such as CA, where most data can only be obtained through third parties, such as the CA WaterInfo portal[14]. Due to the possibility of biases being present in the data, quality assurance should be undertaken prior to analysis and data quality should be considered when interpreting results.

Observed daily precipitation and temperature data were obtained for Naryn (1948-2005), Uzgen (1980-2005) and Sari-Tash (1948-2005) meteorological stations (Figure 1) from the Central Asian Database (CAD)[15]. In the final stages of the PhD, precipitation, temperature and snow depth data were obtained for Lyakhsh and Garm meteorological stations in Tajikistan for the period 2000-2016 from the Tajik Hydromet (the state hydro-meteorological agency). All meteorological data were inspected visually to identify any obvious erroneous measurements or observer bias. Double mass plots were also used between neighbouring stations to identify any step changes indicative of

[14] http://www.cawater-info.net/data_ca
[15] http://www.webpages.uidaho.edu/cae/data/

station relocation. Table 3.3 provides examples of errors identified when quality assuring meteorological data and actions taken as a result. As far as the author is aware, no inhomogeneities are present in the procured meteorological station data (Unger-Shayesteh et al., 2013).

Table 3.3: Examples of errors identified in meteorological data during quality assurance process and actions taken to rectify

Data type	Date	Problem	Action taken
Max temp	12/03/2003	Keystroke error- missing/extra or wrong number	4.3
Max temp	Jun-04	Value given on 31st June	Ignored
Max temp	30/06/2013	Blank cell	-
Max temp	29/02/2016	Leap year missing 29th Feb	-
Max temp	30/09/2005	Missing value symbol not used- cell assigned zero	-
Max temp	26/03/2012	Keystroke error- missing/extra or wrong number	7.1
Mean temp	08/01/2007	Mean outside of range- unsure of reason	-
Min temp	28/05/2012	Keystroke error incorrect decimal place	9.9
Temp	28/08/2001	Transposed min and max temperature	Transposed
Precip	2012	Does '*' mean 0 or missing data	Assigned 0
Precip	2013-16	Blank cells- missing or 0?	Assigned 0
Precip	27/03/2007	Large value duplicates previous day	No action

After the breakup of the Soviet Union the river gauging network declined resulting in the discontinuation of many records (Unger-Shayesteh, 2013). Inhomogeneities were also introduced due to a reduction in gauge maintenance and re-calibration, particularly in headwater catchments. Additionally, procuring recent discharge time series is difficult for headwater reservoirs, as such information is strategically and politically sensitive. Sensitivities stem from the possibility of reservoir inflows and outflows being used by riparian nations to leverage decision making (surrounding the timing of water release/storage). This results in nations typically withholding such information.

Nonetheless, observed mean monthly inflow, storage and outflow were obtained from CAWaterInfo[14] for the years 2001-2010 for Toktogul, Andijan, Kayrakkum and Nurek reservoirs. Historic discharge records were obtained from the Global Runoff Data Centre (GRDC)[16] for the years 1941-1980 for discharges above Toktogul, Andijan and Nurek reservoirs. Missing data (<1%) in the GRDC time series, were filled using the mean monthly value. It is not clear if the same gauge stations were used for the earlier and later records for Toktogul and Andijan, and are known to be different for Nurek. In

---

[16] http://www.bafg.de/GRDC/EN/Home/homepage_node.html

any case, sedimentation and a lack of maintenance of gauges will likely have introduced inhomogeneities between the two data series, potentially reducing comparability. In the final stages of the research, discharge data were obtained for Nurek inflows for the period 1949-2016 (1958-1976 and 1994-1999 missing) from the Tajik Hydromet.

3.3 Gridded products

Despite an extensive search, the only public domain operationally available station-based meteorological data for CA were daily summaries from the National Climatic Data Center[17] and the Global Surface Summary of the Day[18]. For the stations examined, the data were not of sufficient quality for forecasting due to the large percentage of missing values. National hydro-meteorological agencies as well as hydropower companies are known to have access to such data (Apel et al., 2017), but this is not released publicly. Furthermore, the extent to which such data are shared between nations is unclear. Therefore, operational discharge forecasting by entities outside of these organisations/beyond national boundaries must depend of other sources of information. Several gridded precipitation and temperature (Table 2.2 and Table 2.3) products are available for CA, but many are not updated in near real-time (i.e. they have latency >1 week). Such constraints, of being publicly available at operationally useable latencies, severely limit the choice of products.

Recently gridded runoff estimates have been derived from routing land surface/meteorological model outputs (Zaitchik et al., 2010; Emerton et al., 2017). Although in early stages of development, such techniques have potential for discharge estimation in data sparse areas such as CA. Khouakhi (2017) undertook a preliminary evaluation of simulated discharge entering Nurek reservoir following the method of Zaitchik et al. (2010). Two land surface models were used to route the gridded runoff: the Variable Infiltration Capacity and Community Land Model. Inspection of the synthetic discharge data indicate both the timing and magnitude of summer discharges compare poorly with observed data, particularly post-2000 (Figure 3.11). The reason for the unrealistically low synthetic discharge values post-2000 is unclear, but possibly a result of the precipitation inputs. While such synthetic discharge data has significant

---

[17] https://www.ncdc.noaa.gov/
[18] https://data.noaa.gov/dataset/global-surface-summary-of-the-day-gsod

potential in data sparse regions, the poor quality at present led to it not being used in thesis.

Figure 3.11: Comparison of synthetic discharge series derived from gridded runoff with observed discharge (synthetic discharge supplied by Khouakhi, 2017).

3.3.1 Satellite precipitation estimates

Two types of TRMM precipitation estimates were used, allowing a comparison between the 'research grade' product and the real-time 'operational' product. All TRMM products were downloaded from the TOVAS[19] website. The research grade product used is the TRMM Multi Satellite Precipitation Analysis (TMPA) 3B43 (V7) (henceforth referred to as TRMM) monthly precipitation estimates at 0.5° resolution obtained for the period 1998-2010. This product blends information from satellite microwave, radar and infrared instruments from multiple satellites with ground based observations to produce the highest grade product with a latency of 2-3 months (Huffman et al., 2010).

The real-time product is TMPA 3B42 RT (V7) (henceforth referred to as TRMM RT) daily estimates at a 0.5° resolution obtained for the period 2000-2010. This product is available in near real-time (latency of around 2 days), achieved by not including ground based observation correction but instead a climatological correction varying by month

---

[19] https://disc2.nascom.nasa.gov/tovas/

and location (Huffman et al., 2010; Huffman et al., 2017). TRMM RT was aggregated to a monthly time step to align with other data sources. Both products incorporate information from multiple satellites to produce a best estimate of precipitation.

Although the TRMM satellite itself no longer exists, the TRMM product is still produced by incorporating information from other satellites within the TRMM constellation. In the near future it is expected that GPM derived IMERG (the newer version of TRMM) will be available from 1998 onwards, but at present is only available from 2014 onwards. Hence, TRMM is currently the longest running product provided by NASA's precipitation measurement missions. A research grade and operational product were chosen to allow comparison of seasonal discharge forecasts derived from each product. Previous research has shown promising seasonal discharge forecasts using the research grade TRMM product (Dixon and Wilby, 2016). This research aimed to assess whether similar results could be achieved using a potentially operational system.

3.3.2 Re-analysis products

Currently the only operationally available (latency ~2 days) re-analysis product for the region is NCEP/NCAR re-analysis 1 (henceforth referred to as NCEP). The product is also available for a longer time period than TRMM (1948-present), enabling application alongside more widely available historic discharge time series. Therefore, this product was used in the study. Although it is known that NCEP struggles to represent small scale ($<50km^2$) precipitation variations, it does benefit from the denser observational networks in Europe and the Middle East (the main source of moisture to CA) (Schiemann et al., 2008; Unger-Shayesteh et al., 2013). Daily precipitation and temperature data were downloaded at 2.5° resolution for the period 1948-2010 from the SDSM data portal[20]. Data were then aggregated to a monthly time step.

3.4 Large scale climate drivers

The NAO, ENSO and Indian Ocean Dipole (IOD) are evaluated with respect to the regional hydro-climatology in this study. These modes of atmospheric variability were chosen based on previous studies (cited in Chapter 2) as well as on a preliminary

---

[20] http://co-public.lboro.ac.uk/cocwd/SDSM/data.html

analysis of global SSTs in relation to the discharge time series. The Principal Component (PC) based NAO[21] index was preferred to station-based NAO indices because the former provides a more complete characterisation of hemispheric circulation patterns (Hurrell and NCAR Staff, 2016). The Nino 3.4[22] SST based index was selected as the region was strongly correlated with CA discharge time series and is also the ENSO index of choice by forecasters (Barnston et al., 1997). The Hadley Centre Sea Ice and Sea Surface Temperature derived Nino 3.4 index was used as it covers the time period required for the study. It should be noted that this index does not wholly describe ENSO, and may exclude the effect of the western Pacific 'warm pool' identified by Barlow et al. (2002) as important for the ENSO-CA link. The Dipole Mode Index (DMI) describes the IOD, and the SST-based DMI index used here is provided by the KNMI Climate Explorer[23] tool based on data obtained from the Japan Agency for Marine-earth Science and TECnology[24]. All indices were downloaded at a monthly resolution for the period 1940-2016.

3.5   Evaluation of gridded data products

To evaluate to relative qualities of the gridded precipitation and temperature products used in the study, comparisons were made with observed data. In each case, the coincident TRMM (0.5°) or NCEP (2.5°) cell overlaying the gauge was used. Time-series were plotted at each gauge site allowing a visual inspection of the data, highlighting any periods where products deviated from observed values. Monthly, seasonal and annual precipitation totals were compared to assess any seasonal/annual over/under estimations. Correlation analysis allowed the strength of relationships between gridded and observed precipitation and temperature to be assessed. These procedures, along with previous literature identified in section 2.3.1, provide a deeper understanding of the gridded precipitation and temperature products relative quality. This was of particular importance when assessing the performance of regression based seasonal forecasting models, as particular problems with input variables may help explain issues with the final models.

---

[21] https://climatedataguide.ucar.edu/climate-data/hurrell-north-atlantic-oscillation-nao-index-pc-based
[22] https://climatedataguide.ucar.edu/climate-data/nino-sst-indices-nino-12-3-34-4-oni-and-tni
[23] https://climexp.knmi.nl/
[24] http://www.jamstec.go.jp/e/

TRMM RT (available from March 2000 onwards) overestimates gauge precipitation for every year in the record (Figure 3.12). At Naryn, TRMM RT overestimates annual gauge precipitation by 283% on average, compared with 2% and 19% overestimations by TRMM and NCEP respectively. TRMM RT performs better at Uzgen than Naryn and Sari-Tash, overestimating average annual precipitation by 44%, with an average overestimation of 12% for the years 2003-05. Overestimation by TRMM RT is not consistent throughout the year (Figure 3.12). In particular, May and October seem to be problematic months, with overestimation exceeding 170% for all basins. This could be due to uncertainty in the estimation of snowfall from satellites during these months or the fact that increased precipitation occurs during these months (Gebremichael et al., 2010; Yong et al., 2012). The good agreement between TRMM and gauge precipitation, therefore, appears to be achieved via gauge-correction of the data, rather than satellite skill. NCEP performs much more favourably than TRMM RT, but still struggles to capture the annual regime. For example, underestimating April (-41%) whilst overestimating August (181%) observed precipitation at Naryn. On average, TRMM performs best of all products evaluated when compared with gauge observed precipitation for both annual regime and annual totals.

Another potential cause of satellite overestimation suggested by Dixon and Wilby (2016) was local heavy precipitation events under southerly monsoon airflows. This was reasoned because of the timing of overestimation by TRMM at Naryn (July, August, September). TRMM RT consistently over-predicts during these months for all three sites (Figure 3.12). This adds weight to the argument that high intensity local precipitation events may be missed by the gauge point measurement but detected by the 0.5° TRMM cell (Bowman, 2005; Bothe et al., 2012).

Although NCEP data have been used extensively for climate studies, few seem to have used the product in Central Asia. It has been suggested this is because of the spatial resolution being too coarse to capture the detailed topographic influence on climatic conditions in the region (Schiemann et al., 2008; Hu et al., 2016). At Naryn, NCEP overestimates precipitation during July and especially August (Figure 3.12). The relatively large cell size of NCEP could be capturing locally heavy precipitation events under southerly monsoon airflows that the gauge does not detect. For Sari-Tash, an underestimation during summer, and a marked overestimation during winter is present

compared to observed precipitation. This may be a result of the stations close proximity to the Alay Mountains to the west (rising to over 3,800m within 1km). Such fine scale topographic variation cannot be accounted for by NCEP due to the coarse (2.5°) spatial resolution of the re-analysis products.

Monthly TRMM precipitation estimates are strongly correlated with gauge measurements both at Naryn ($r = 0.97$) and Andijan ($r = 0.92$), but less so for Sari-Tash ($r = 0.63$) (Figure 3.13 and Table 3.4). NCEP precipitation displays consistently stronger correlation than TRMM RT for all basins. All three products show relatively strong correlations at Andijan compared with the Sari-Tash gauge. TRMM RT shows severe overestimation (exceeding 450%) at all three basins. These biases typically occur in either April-June or September-November – the periods of precipitation maxima at all three sites. It is possible that the overestimation is a caused by precipitation falling as both snow and rain during this period as well as the large precipitation totals. Satellite precipitation estimates are known to perform poorly during snowfall, and TRMM RT is therefore attributed a constant precipitation rate (Gabremichael, et al., 2010; Huffman and Bolvin, 2017). However, if the precipitation phase is changeable, it is possible that the constant rate is being applied inconsistently during some snowfalls resulting in spurious values. The TRMM product does not suffer from these large overestimation issues, likely rectified during the gauge adjustment process.

NCEP temperature shows consistently strong correlations to measured temperature for all stations investigated ($r \geq 0.9$) (not shown). At Naryn, and to a lesser extent Sari-Tash, a summer/winter break point in the relationship between station and NCEP temperature is evident (Figure 3.14). The cause of this deviation is unclear, with minimal information available on meterological station location and site conditions. It is plausible that local summer temperatures at the station could be depressed by proximity to vegetation or the river, yet no photographs of meteorological stations could be found to corroborate either. As the effect is seen at both Naryn and Sari-Tash, the problem could derive from the NCEP re-analysis. The different albedo's during winter and summer due to snow cover at the meteorological stations could plausably cause different relationships between NCEP and local temperature during summer and winter.

Figure 3.12: Annual precipitation totals (left) and monthly mean precipitation (right) for Naryn (upper), Uzgen (centre) and Sari-Tash (lower) at gauge locations for the years 1998-2005 (2000-2005 for TRMM RT). See Figure 3.1 for gauge locations

Figure 3.13: Monthly gauge observed precipitation plotted against coincident cell monthly gridded precipitation for the period 03/2000-12/2005. Dashed line shows 1:1 relationship.

Table 3.4: Spearman rank correlation coefficients between monthly precipitation estimates (from TRMM, TRMM RT, and NCEP) and monthly gauge observed precipitation at meteorological stations. Values in bold are significant at *p* = 0.05.

	TRMM	TRMM_RT	NCEP
Naryn	**0.97**	**0.35**	**0.50**
Uzgen	**0.92**	**0.45**	**0.82**
Sari-Tash	**0.63**	0.21	0.32

Figure 3.14: Observed monthly temperature at Naryn plotted against the coincident cell NCEP re-analysis temperature for the years 1948-2000.

Despite this, the high correlation at all stations suggests NCEP can capture the annual regime of temperature sufficiently well, which in combination with its short latency (~2 days) allows it to be used in place of station temperature when forecasting reservoir inflows.

3.6   Summary

The data procured for the study were chosen in accord with previous seasonal forecasting efforts in the region (section 2.5). In order to address the overall aim and

objectives of this research it is paramount that suitable data are collated. Key qualities for operational forecasting of discharge are near real-time, and in this instance, public domain accessibility. Whilst hydro-meteorological agencies in the region have near real-time access to observational networks, public domain information could provide several benefits. For example, improved spatial coverage of climate monitoring and avoiding issues associated with data sharing between riparian states. With this potential in mind, undertaking an exploratory study of the use of public domain datasets is justified.

Meteorological station precipitation and temperature data have been procured in each of the four study basins, allowing quality assessment of gridded products. Observed discharge data, required for model building and validation, have also been obtained for the four study basins. It is clear from the analysis presented in Section 3.5 that the gridded products to be used for seasonal discharge forecasting do not always align well with observed values of precipitation and temperature at monthly resolution. Monthly TRMM RT overestimates precipitation, particularly during spring and autumn, whilst also showing weakest correlation with observed precipitation. However, all three monthly precipitation products (at Naryn and Uzgen) and temperature (all stations) show significant correlations with monthly observed data. Such results provide promise that these products can be used for seasonal discharge forecasting in the region.

Sourcing high quality hydro-climatological data can be challenging in CA and a significant obstacle to research. The observed hydrological and meteorological data collated here derive from several sources which, along with lack of maintenance, observer bias and data entry issues, could introduce error into observed time-series. In data sparse regions where hydro-climatological information may be a politically and economically sensitive issue, it is necessary to be creative about the use of unconventional data types and estimates of surface observations. Table 3.5 provides an overview of all the data sets used in the following research. The next chapter outlines the methodological approaches to address the aim and objectives of the thesis using the information presented in this chapter. Model choice is justified following which a description of the development, calibration and validation of the statistical and climate mode base forecasting approaches id provided.

Table 3.5: Summary of data used in the study

Parameter	Data set	Temporal coverage	Temporal resolution	Spatial resolution	Source
Precipitation and Temperature	Central Asian Database	(1930-1980) - 2005	Daily/ Monthly	Station observed	Central Asian Database[15]
Precipitation, Temperature and Snow height	Tajikistan hydro-meteorological agency	2000-2016 (1958-76 and 1994-99 missing)	Daily	Station observed	Tajikistan hydro-meteorological agency
Discharge	CA Water Info	2001-2010	Decadal/ Monthly	Station observed	CA Water Info[14]
Discharge	GRDC	1941-80	Monthly	Station observed	GRDC[16]
Discharge	Tajikistan hydro-meteorological agency	1949-2016	Daily	Station observed	Tajikistan hydro-meteorological agency
Discharge	Synthetic discharge	1979-2016	Daily	N/A	Khouakhi, 2017
Precipitation	TRMM TMPA 3B43 (V7)	1998-2017	Monthly	0.5°	TRMM TOVAS[19]
Precipitation	TRMM TMPA 3B42 Real Time (V7)	2000-2017	Daily	0.5°	TRMM TOVAS[19]
Precipitation and Temperature	NCEP/NCAR re-analysis 1	1948-2017	Daily	2.5°	SDSM download portal[20]
North Atlantic Oscillation	Hurrell PC NAO Index	1899-2017	Monthly	N/A	Climate Data Guide[21]
El Niño Southern Oscillation	Niño 3.4	1870-2017	Monthly	N/A	Climate Data Guide[22]
Indian Ocean Dipole	Dipole Mode Index	1870-2017	Monthly	N/A	KNMI Climate Explorer[23]

4 Methodologies

The previous chapter introduced the study sites and data used in the research. What follows is a description of the methodological approaches used to address the objectives of the study. Seasonal forecasting model choice is first justified; following which the development and validation of the statistical model is outlined. A second approach to seasonal forecasting is then described, using modes of climate variability to condition mean summer discharge. The chapter concludes with a summary of the main strengths and weaknesses of the methodologies adopted.

4.1 Model choice

Section 2.4.1 reviewed available seasonal forecasting model types, including dynamical, analogue and empirical approaches. The large data demand and computing power of dynamical models reduce their usability in data sparse, low capacity regions. Conceptual water balance models have been used in the region for climate change assessment (e.g. Wilby et al., 2011a), but these require meteorological data outlooks to enable seasonal discharge forecasting. The large data requirements, difficulty with calibration (especially in regions unknown to the modeller) as well as lack of transparency in model structure resulted in dynamical/conceptual approaches being deemed unsuitable. The minimal data requirements and relatively strong autocorrelation of the discharge data make an auto-regressive approach (such as ARIMA) more appealing. However, previous literature and current forecasting practices of Tajik Hydromet (regression of winter precipitation to forecast mean summer discharge) suggest that winter precipitation conveys skill to forecast following summer discharge (Schar et al., 2004; Barlow and Tippett, 2008; Dixon and Wilby, 2016; Apel et al., 2017). This skill would be ignored by a simple lag-1 auto-regressive model, thus such an approach was unsuitable.

A preliminary study was undertaken to evaluate an Artificial Neural Network (ANN) approach. An ANN model was piloted by Dawson (pers. comm.) to forecast inflows to Toktogul reservoir allowing comparison with Dixon and Wilby (2016). Results showed that similar skill levels could be achieved by the ANN model compared with a multiple

linear regression approach. However, ANN's are non-transparent black-box models resulting in internal weights being difficult to interpret (Abrahart et al. (2012). This, along with increased computing demand compared with multiple linear regression meant the approach was not preferred.

After reviewing available modelling options, ordinary least squares multiple linear regression was considered to be the most suitable approach for this study. Regression was chosen because the technique is:

- Transparent
  - Allowing ease of input data and operator error tracking
- Tractable
  - Given the data and skill capital of the region
- Transferable
  - The technique is already recognised and applied by practitioners easing application across CA

However, regression is not without weaknesses since a long enough calibration data period is required to capture catchment characteristics (Magar and Jothiprakash, 2011). It is also, in theory, unsuitable to be applied to other catchments without re-calibration (Wagener et al., 2004). Despite these issues, the advantages of regression as well as its previous successful implementation in nearby regions (e.g. Archer and Fowler, 2008; Dixon and Wilby, 2016) were such that it was considered most appropriate.

4.2   Statistical forecasting model development

4.2.1   Predictor selection

For every cell in each basin concurrent and lagged (0-3 months) TRMM and TRMM RT precipitation as well as NCEP precipitation and temperature were correlated with reservoir inflows. Correlations were assessed via the non-parametric Spearman's rank-order correlation, therefore not requiring predictors to be normally distributed. This was undertaken to identify the cell(s) and area(s) of each basin and lag interval that potentially yield greatest predictability of inflows. Then, following Dixon and Wilby (2016), simple linear regression relationships were explored between runoff in the summer melt season (April-September) as a function of individual predictors in the preceding winter (October-March). This split-year approach has been applied to other runoff records in Central Asia and informs the preliminary selection of candidate

predictor variables for sub-seasonal forecast models (e.g., Schär et al., 2004; Pal et al., 2013). Following this, predictor variables were correlated with discharge during the period October 2001 to December 2010 at moving average periods 1-6 months and lag intervals 0-4 months for each basin. This allowed an assessment of suitability of candidate predictors for inclusion in multiple linear regression models. Available monthly variables were lagged and time averaged TRMM, TRMM RT and NCEP precipitation as well as NCEP temperature at both individual cell (identified in step 1 of predictor selection) and basin average spatial scales. Furthermore, antecedent flow and a dummy variable for month were used to capture the average annual cycle (Table 4.1). Individual TRMM and NCEP cells were used to allow contributing area(s) with greatest potential predictability to be captured within the models. Dummy variable regression weights are equal to the flow anomaly with respect to a reference month, in this case December.

4.2.2 Model construction and calibration

Regression modelling proceeded as in Dixon and Wilby (2016), first building the model with a small set of statistically significant predictors before model skill was assessed via a k-fold (k = 9) cross validation. One change to the method was an extension of the lead times assessed from 1-3 to 1-4 months. Two types of model were created: 'research' grade and an 'operational' version (Table 4.2). Research grade models include all available predictors, whereas operational versions used only those predictors that would be available in (near) real-time. This allowed assessment of the potential of Dixon and Wilby's (2016) method to be implemented operationally.

Regressions were undertaken by first including the monthly dummy variable, then running a stepwise selection procedure with all other predictors. This process involved variables being added one at a time beginning with the predictor with the lowest probability of F (prob(F)). The prob(F) is the probability that the null hypothesis (i.e. that a predictors regression coefficient is 0) is true. A predictor is only retained by the model if it complies with the set value of prob(F) (F-to-enter). After each step, any predictors already in the regression equation with a prob(F) larger than the set value (F-to-remove) are removed. The stepwise procedure terminates when no more predictors are eligible for inclusion or exclusion in the regression equation. Very little information could be

found to guide the selection of F-to-enter/remove values. Relatively stringent parameters of F-to-enter = 0.01 and F-to-remove = 0.05 were selected to reduce the likelihood of over-fitting the model to the training dataset (Wilks, 2011). This ensured all predictors in the final regression models had a prob(F) < 0.05. Following termination of the stepwise procedure the sequence of variables entered/removed was reviewed to ensure the variables included/excluded made sense physically and that the greatest adjusted R^2 was achieved by the final model. Several assumptions are associated with linear multiple regression, including predictors having a linear relationship to the predictand as well as independence, normality and homoscedasticity of the errors (Wilks, 2011). The performance of regression models with respect to these assumptions is assessed in section 4.2.4.

Stratified models were created for forecasting discharge for only summer months (April-September). This stratification procedure was used due to the strong seasonality of discharge and is comparable to the approach adopted in basins with similar regimes (e.g. Archer and Fowler, 2008; Barlow and Tippett, 2008). These summer stratification models were only produced as operational versions, i.e. only including predictors that would be available in near real-time. Combinations of predictor variable, moving average period and lag interval were evaluated to forecast only summer monthly inflows. The same stepwise-regression procedure was used as for the research grade models. A summary of the models and their respective inputs and target periods is provided in Table 4.2.

Table 4.1: Independent variables used in regression models (note: Andijan and Toktogul outflow were only used for Kayrakkum inflow modelling)

Notation	Description
$TRMM\_Av_{t,n}$	0.5° TRMM upstream catchment area average precipitation estimate
$TRMM\_Op_{t,n}$	0.5° TRMM optimum cell precipitation estimate
$TRMMRT\_Av_{t,n}$	0.5° TRMM real-time upstream catchment area average precipitation estimate
$TRMMRT\_Op_{t,n}$	0.5° TRMM real-time optimum cell precipitation estimate
$NCEPP\_Av_{t,n}$	2.5° NCEP re-analysis upstream catchment area average precipitation estimate
$NCEPP\_Op_{t,n}$	2.5° NCEP re-analysis optimum cell precipitation estimate
$NCEPT\_Av_{t,n}$	2.5° NCEP re-analysis upstream catchment area average temperature estimate
$NCEPT\_Op_{t,n}$	2.5° NCEP re-analysis optimum cell temperature estimate
$Q_{t,n}$	Antecedent discharge measured at reservoir inflow
$AND_{t,n}$	Observed outflow from Andijan reservoir
$TOK_{t,n}$	Observed outflow from Toktogul reservoir
t	Variable lag interval (t months)
n	Variable averaging period (n previous months)
M	Dummy variable for calendar month

Table 4.2: Summary of models constructed

Model type	Available inputs	Forecast target period
Research	All variables in Table 4.1	All months for the period October 2001-December 2010
Operational	All variables in Table 4.1 excluding $TRMM\_Av_{t,n}$ and $TRMM\_Op_{t,n}$	All months for the period October 2001-December 2010
Summer-only	All variables in Table 4.1 excluding $TRMM\_Av_{t,n}$ and $TRMM\_Op_{t,n}$	Months April-September for the period April 2002-September 2010

4.2.3 Model validation

Regression model skill was assessed using a leave one out k-fold cross-validation by which monthly discharge for individual years was predicted using models built on all other years of data. For instance, flows for the year January 2010-December 2010 would be predicted by a model calibrated on data for October 2001-December 2009. Year-by-year a full series of predicted flows was built enabling validation against data not used in model calibration. Available reservoir inflow records permitted cross-validation of 9 year-long segments of data (the first segment being 15 months long), each with their own regression parameter sets. This provides a more stringent test of model skill than measures of calibration fit to the whole data. Note that, however, an operational version of the model would be fit to the selected predictor set using the entire record then recalibrated periodically as more data become available.

The skill of all model predictions was benchmarked with against the Zero Order Forecast (ZOF). This is the amount of explained variance that can be obtained from the simplest possible model – in this case the long-term monthly mean flow. The long term mean monthly flow was chosen rather than a persistence forecast (i.e. the current months discharge will be next months) due to it harbouring greater skill and so being a more stringent benchmark. For comparability, identical blocks of data were used for estimating the long-term mean as those entered into the nine cross-validated regression models. Summer only models were benchmarked against the ZOF as well as summer months extracted from the full-year forecasts to assess any improvements yielded by summer stratification relative to the full year operational models.

The HydroTest tool of Dawson et al. (2007) was used to derive five metrics of model forecast skill: the amount of explained variance (R^2); Root Mean Squared Error (RMSE); Akaike Information Criteria (AIC); Nash-Sutcliffe Coefficient (NSC); and the Mean Absolute Relative Error (MARE). A review of each skill metric is provided in Table 4.3. These were chosen accounting for the aim of the study (scientific, i.e. evaluating strengths and weaknesses of models), the nature of the predictand (continuous) as well as the attributes of the forecast to be verified (bias: correspondence between mean forecast and mean observation; accuracy: average correspondence between individual pairs of observations and forecasts and skill: accuracy of forecasts relative to accuracy of forecast benchmark) (Wilson, 2004). These metrics allow forecasts to be evaluated

for low (MARE) as well as high (RMSE) flows, accounting for model complexity and calibration period (AIC) and the fit of the forecast to the regression line (R2 and NSC).

Table 4.3: Relative strengths and weaknesses of selected skill metrics.

Skill Metric	Definition	Strengths	Weaknesses
R^2	Squared ratio of the combined dispersion of two series to the total dispersion of the observed and modelled Series.	Describes the proportion of the total statistical variance in the observed dataset that can be explained by the model.	Insensitive to additive and proportional differences. Oversensitive to outliers and thus biased towards a consideration of extreme events.
RMSE	The level of overall agreement in real units between the observed and modelled datasets.	Sensitive to forecasting errors that occur at higher magnitudes.	Insensitive to forecasting errors at lower magnitudes. Sensitive to an occasional large error.
AIC	RMSE adjusted according to the number of free parameters and data points.	Attempts to account for model complexities. Attempts to find the minimal model that best explains the dataset, discouraging overfitting.	Does not use real units. Same issues as RMSE.
NSC	One minus the ratio of sum square error to the statistical variance of the observed dataset about the mean of the observed dataset.	Sensitive to differences in the observed and modelled means and variances (unlike R^2).	Similar weaknesses as R^2, being insensitive to additive and proportional differences and oversensitive to outliers.
MARE	Mean of the absolute error made relative to the observed record.	Sensitive to forecasting errors that occur at lower magnitudes.	Insensitive to larger errors typically occurring at high magnitudes.

4.2.4 Assumptions of regression

Multiple linear regression analysis depends on several assumptions. The compliance of regression models created within section 6.1 with these assumptions was tested using the methods and acceptance criteria provided in Table 4.4. Both visual inspections and statistical metrics were used to assess model compliance. Several models were tested

for a range of lead times at each reservoir. For brevity, representative models were chosen to demonstrate the overall compliance of models used (Table 4.5).

Stratified models comply with assumptions better than full year models, with linear relationships between dependent and independent predictors. For full year models, not all predictors were linearly correlated to inflows. To assess the impact of this non-compliance on Toktogul inflow forecasts at a one month lead, the model was re-calculated after transforming both the dependant variable (Toktogul inflows) and the other independent variable (antecedent flow) using the natural logarithm (as used in previous studies, e.g. Svensson et al., 2015; Mendoza et al., 2017). This resulted in an improved R^2_{adj} (not cross validated) of 0.94 compared with 0.89 before transforming the data. Applying the transformation results in linear relationships between all dependent and independent variables, residuals being normally distributed and improved homoscedasticity of residuals. If the statistical forecasting approach were to be operationalised, such a transformation should be applied to ensure the model conforms to key assumptions and optimal performance is achieved using the predictors available. The violation of assumptions shown here highlights the dangers of using a stepwise regression procedure, with the risk of it becoming a 'black box' model. Due to this, expert judgement should always make the final decision regarding predictor variables and assumption compliance should always be evaluated.

Table 4.4: Multiple linear regression assumptions and criteria used to assess model compliance.

Assumption	Assessment technique	Acceptance criteria
Linearity between the dependant and independent variables (both individually and collectively)	Scatterplot	Visual inspection
No multicollinearity of independent variables	Variance Inflation Factor (VIF)	VIF <10
Normal distribution of residuals	Shapiro-Wilk test	*p*-value greater than 0.05
Homoscedasticity through time	Scatterplot of residuals against time	Visual inspection
Homoscedasticity through predicted flow values	Breusch-Pagan test	*p*-value greater than 0.05
No autocorrelation present in residuals	Durbin-Watson statistic	Test statistic between 1.5 and 2.5

Table 4.5: Selected models compliance with multiplelinear regression assumptions outlined in Error! Reference source not found.**.**

Assumption	Toktogul Q1 full year	Toktogul Q1 stratified
Linearity	Q = PASS NCEP T = FAIL Combination = PASS	NCEP_P_Op = PASS NCEP_P_Av = PASS TRMMRT_Op = PASS Combination = PASS
No multicollinearity	Q = PASS NCEP T = FAIL	NCEP_P_Op = PASS NCEP_P_Av = FAIL TRMMRT_Op = PASS
Normal distribution of residuals	FAIL	PASS
Homoscedasticity through time	PASS	PASS
Homoscedasticity through predicted flow values	PASS	PASS
No autocorrelation present in residuals	PASS	PASS

4.3   Climate mode based seasonal forecasting

4.3.1   Correlation analysis

The climate mode seasonal forecasting approach followed the methodology presented in Dixon and Wilby (under revision). First, climate modes were selected based on previously recorded relationships with CA precipitation and/or river flow as well as a preliminary global analysis to identify ocean basins (using SST's) with potential predictability of inflows to Toktogul, Andijan and Nurek reservoirs. Kayrakkum reservoir inflows were not forecast in this way. This was due to discharge being determined by upstream river regulation rather than snow and glacier melt. The ESRL correlation tool[25] was used to correlate monthly NAO, Niño 3.4 and DMI indices with NCEP re-analysis precipitation and temperature across CA for each month during the period 1950-2014, over lagged intervals (index leading) 0 to 3 months. Individual monthly, bi-monthly and winter mean climate mode indices were then correlated with following mean summer inflows into Toktogul, Andijan and Nurek reservoirs. Summer flows were selected as these contribute 77% to 83% of mean annual inflow. Two periods were used, 1941-80 and 2001-2010 to assess the stationarity of correlations.

4.3.2   Composite forecasting

Average summer inflows were pooled conditional on the tercile (i.e. <33%, 33-66%, >66%) of prior winter mean indices. The period 1941-1980 was used to minimize possible non-homogeneity due to discharge data and/or anthropogenic climate change in the region. This tercile approach enables evaluation of equal size samples of summer flow conditional on prior index (negative, neutral or positive). Statistical significance of differences between pooled samples was assessed using two tests. First, the Kruskal-Wallis (KW) test assesses the whether the ranked values of multiple distributions derive from populations with the same distribution. The test determines if at least two groups of conditioned flows are significantly different, but does not state which specific groups. Second, the Kolmogorov-Smirnov (KS) test assesses whether two independent samples have been drawn from the same population/from identical populations. The test produces a *D* statistic which is a measure of the maximum difference between two cumulative distributions. Both tests do not make assumptions regarding the distribution

---

[25] https://www.esrl.noaa.gov/psd/data/correlation/

of data. Box-and-whisker plots were used to show expected flows following negative/neutral/positive index phases. Concurrent summer air temperature terciles were also used (as a proxy for snowmelt contributions) to further condition summer inflows pooled by winter index.

Further analysis focused on Nurek inflows as these showed greatest potential for conditioning on prior climate mode indices and concurrent temperature. The stationarity of the Niño 3.4-Nurek inflow relationship was assessed followed by an example of how a forecast based on this relationship could be issued. The procurement of Additional long term, up to date discharge data for Nurek allowed stationarity to be assessed by undertaking composite analysis for the period 1981-2016 (1993-1999 missing) as well as 1941-1980. Up-to-date discharge data was procured for Komsomolabad (downstream of Garm) allowing a 'blended' long term time series (1941-2016 (1993-1999 missing) to be constructed. This blended series was constructed for Komsomolabad by regressing mean summer (April-September) flows against mean summer Garm flows using overlapping periods of data (1949-1957, 1977-85, 1988 and 1990). The regression model explained 68% of the variance in mean summer flow at Komsomolabad. Mean summer flows were used to build the regression equation instead of monthly flows to limit autocorrelation of errors ($r=0.34$, insignificant at $p=0.1$). The flow data used to build the model was also stationary over the period used. The final blended flow series was created using observed values where possible, filling any gaps in the series with predicted values from the regression model. No data was available from either gauge for the period 1993-1999. The same procedure was applied to build a long-term dataset for Sari-Tash temperature using its relationship with Lyakhsh allowing sensitivity testing of inflows conditioned on concurrent temperature.

Using October-March average index, a forecast cannot be issued before early April. After discussion at a workshop attended by local stakeholders, including members of Tajik Hydromet and Barqi Tojik, it was apparent an early outlook forecast (in addition to Tajik Hydromet's April issued forecast) was desired. Therefore, average summer inflows to Nurek were conditioned on mean November-December Niño 3.4, allowing a forecast to be issued in January. November-December Niño 3.4 was used due to its strong correlation with mean summer inflows during both 1941-1980 and 2001-2010. Forecasts are presented using likelihood tables, giving the likelihood (%) of flows falling

between various anomaly ranges. Forecast skill for both issue times was assessed via the Heidke Hit Proportion (HHP), defined as the proportion of forecast cases in which the forecast category assigned the highest probability is later observed (International Research Institute, 2013). A score of 0.33 would be expected by chance alone; therefore, a greater value indicates additional skill is provided by the model. The time series was split in half, allowing cross validated HHP scores to be calculated for 1941-1974 and 1975-2016. Sensitivity testing of inflows conditioned on prior Niño 3.4 and temperature was undertaken by varying the flow anomaly baseline, the period used to derive tercile boundaries and the met station used to derive temperature.

4.4   Transferability of methodologies

To assess the transferability of the methodologies piloted in the study, both regression and climate mode based forecasting techniques were implemented on the Ouergha River, in the headwaters of the Sebou basin, Morocco (Figure 4.1). The M'Jara gauge is located 5km downstream of the subsequently built Al Wahda dam and reservoir, the largest in Morocco (commissioned in 1997, transferability will be tested for the period 1952-89). The location of the gauge at a now significant headwater reservoir, used for flood control, irrigation and hydropower production, means accurate seasonal forecasting could help reduce drought impacts as well as prevent over-topping as seen during 2009-2010 (Abdelbasset et al., 2015). Therefore, although data are only available pre-dam building, the basin was thought suitable for assessing the transferability of approaches.

The hydro-climate of the area is very different to that of CA. Both precipitation and discharge peak during January-December and are virtually nil during July-August. The basin is also located close to the coast inland from the Atlantic at a lower elevation (maximum of 2440m) than the CA basins. The basin area above the gauging stations is ~6000km^2. Generally, precipitation falls as rain, but snow is known to fall over the mountainous region in the north-east of the catchment. Such a contrasting climate will provide a stern test of the approaches

75

Figure 4.1: Location of M'Jara basin within Morocco

Monthly discharge data was procured at M'Jara for the period 1952-1989 from GRDC[16], with missing values (<3%) being filled with monthly means. For regression modelling, monthly NCEP precipitation and temperature (basin average and optimal cell) are used as predictors along with previous discharge. The testing period is prior to TRMM satellite precipitation availability, therefore it could not be used. The same regression modelling procedure was used as in CA, however only operational models (predictors available in near real-time) were constructed, calibrated and validated (k-fold validation excluding 5 years at a time) for lead times 1-3 months. No seasonal stratification was applied.

Figure 4.2: Mean discharge regime of the Ouergha River at M'Jara for the years 1952-1989.

Hurrell (1995) (principal component) NAO and HadISST derived Niño 3.4 were selected as potential climate modes to be used for conditioning discharge. Climate mode-based forecasting was undertaken as before, first correlating precipitation and temperature with NAO and Niño 3.4 before conditioning discharge according to prior index phase. Two target discharge periods were investigated: December-January mean for dam safety; and February-April mean for reservoir filling. Potential management actions with sufficient lead time forecasts might include reservoir drawdown to reduce potential of overtopping or reduced winter release for hydropower generation if low inflows are expected during February-April.

4.5   Overview of approach

An overview of the methodological approach used in the study is provided in Figure 4.3. Observed hydro-climatic data are difficult to procure for CA, which has led to much information being gathered third hand. This cannot be avoided in a region in which hydro-climatic data are so politically sensitive, however quality assessment has been undertaken and data integrity should be kept in mind when drawing conclusions.

The relative merits of modelling approaches were outlined in section 4.1. Using a multiple linear regression approach requires several assumptions to be met and the extent to which models used here comply with these was evaluated in section 4.2.4. Regression models will be validated using a leave one out k-fold cross validation (for CA models, leave five years out for Morocco). It is known that data occurring after a forecast is issued can unfairly advantage a model when using a leave one out cross validation compared to real time application (Robertson, 2016). However, with such a short record length (10 years) omitting several years of data trailing each validation period would leave too little data on which to calibrate the model. Therefore, the potential increased skill of cross validated models compared to real world application should be acknowledged when evaluating model performance. The monthly mean discharge was used as a measure of ZOF rather than a persistence forecast. This was chosen as mean monthly discharge provided a more stringent test with which to benchmark model skill. The ZOF is the simplest possible model, so provides a useful test of the value-added by more sophisticated approaches.

77

Monthly precipitation and temperature series were correlated with concurrent and preceding monthly climate mode index values. This single month approach was utilised as it allowed greatest resolution of climate mode index skill to be depicted. A bi-monthly approach was also evaluated and whilst slightly stronger correlations were found, there was less precision about which exact periods provide greatest skill. Boundaries for conditioning of discharge by prior index phase were set using 0.33 and 0.66 percentiles, rather than prescribing fixed boundaries for negative/neutral/positive phases. This allows roughly even numbers of discharge seasons for each phase but does mean events could be classed as positive/negative during relatively weak events. Furthermore, even sized terciles can result in large variations of event strength *within* each phase.

```
                        Data procurement
                       ↙              ↘
        Evaluate gridded products      Indices correlated with CA
                 ↓                     precipitation and temperature
        Winter predictors correlated            ↓
        with following summer flow     Prior indices correlated with
                 ↓                     following summer mean discharge
        Predictor variables correlated          ↓
        with discharge at various       Summer inflows pooled conditional
        averaging periods and lag       on the tercile of prior winter SSTs
        intervals                               ↓
                 ↓                      Assess statistical significance of
        Construct regression models     differences between pooled samples
        using stepwise procedure                ↓
                 ↓                      Tables and box and whisker plots
        Validate models using k-fold    used to show expected flows following
        cross validation                negative/neutral/positive index phases
                 ↓                              ↓
        Skill assessed via several      Skill assessed using Heidke Hit
        metrics and benchmarked         Proportion
        against ZOF                     ↘
                 ↘              ↙
           Explore transferability of approach to a
           basin in Morocco
```

Figure 4.3: Conceptual diagram of methodological approach

4.6   Summary

This chapter has described the methodologies used to address the aim and objectives of the study. Justification was provided for the use of multiple linear regression. It is acknowledged that other approaches could have been applied, but regression was

judged most suitable due to it being transparent, tractable and transferable. Detail has been provided regarding both regression and climate mode-based forecasting approaches. Information regarding the M'Jara basin in Morocco, the chosen site for testing transferability of approaches, helped contextualise the site and drew comparison with CA. Section 4.5 offered a review of the methodologies, discussing associated strengths and weaknesses. The following chapter will present the results of analyses executed using the above outlined techniques.

5 Results

The results of the thesis are presented in this chapter produced using the methodologies introduced in Chapter 4. First, potential for seasonal forecasting inflows to four headwater reservoirs in Central Asia is assessed. Forecasting was undertaken at a monthly time-step for the period 2001-2010 using a statistical (regression) model for lead times 1-4 months. Results are presented in three stages through predictor selection, model construction and calibration, then validation. Second, forecasts are piloted for mean summer flow using climate modes as predictors. Summer inflows to Toktogul, Andijan and Nurek reservoirs for the period 1941-80 were conditioned on prior winter Niño 3.4, NAO and DMI phases (negative/neutral/positive) to assess each index's discriminative power. A method to operationalise the approach is then piloted for Nurek reservoir. The transferability of both approaches is subsequently assessed via application to a river basin in Morocco. The chapter concludes with a summary of headline findings.

5.1 Statistical forecasting model development

5.1.1 Predictor selection

The amount of variance in monthly inflows explained by overlaying cells of TRMM (Figure 5.1) and TRMM RT (Figure 5.2) precipitation as well as NCEP precipitation (Figure 5.3) and temperature (Figure Apx 1) for lead times 0-4 months were explored, extending Dixon and Wilby's (2016) analyses. Correlations were tested via Spearman's rho, with significant values taken at $p = 0.05$. Due to the autocorrelated nature of the inflow data, the 'effective sample size' was adjusted downwards following Wilks (2011). Critical rho values for each basin, accounting for autocorrelation, can be found in Table 5.1. This was undertaken as a preliminary analysis to explore predictor potential, with all months utilised to maximise the small sample size. It should be noted, however, that these correlation surfaces incorporate different hydrological conditions throughout the year. This means winter snowpack accumulation, summer ablation period and groundwater dominated flows during winter are mixed in this preliminary analysis.

Inflows to Toktogul and Andijan showed similar correlation patterns with TRMM and TRMM RT precipitation products at all lead times (Figure 5.1, Figure 5.2 and Figure

5.3). Cells with greatest explained variance at zero lead include central and eastern areas of each basin, over the Tien Shan Mountains. As the lag interval increased, the zone of greatest predictability migrated towards the north-west. In contrast to TRMM and TRMM RT, NCEP precipitation provided greatest predictability of inflows to Andijan at a three-month lead. NCEP precipitation is increasingly positively correlated with Andijan inflows as lead time increases, contrasting the relationships of TRMM and TRMM RT (Figure 5.3).

Figure 5.1: Monthly observed inflow into Toktogul, Andijan, Kayrakkum and Nurek reservoirs correlated with monthly TRMM estimated precipitation for the years 2001-2010 for concurrent (left) and precipitation leading (right) discharge by 1, 2, and 3 months. Note that a 4 month lead was also undertaken but not shown for brevity.

Kayrakkum exhibited the opposite relationship to both Toktogul and Andijan. At zero lead, significant negative correlations were observed over the mountainous cells in the east. The transition from negative to positive correlation pattern occurred at the location of Toktogul reservoir. TRMM and NCEP precipitation products were negatively correlated with Nurek inflows concurrently, transitioning to positive correlations with increasing lead times. TRMM RT showed an opposing relationship, with greater negative correlations occurring at a three-month lead.

Although similar patterns of correlation were found, TRMM RT explained consistently less variation in inflows than TRMM for all basins, with fewer cells exhibiting statistically significant correlations at each lead time. NCEP showed comparable correlations at each lead time for each basin as TRMM, with similar areas of each basin significant at each lead (with the exception of Andijan). The different spatial resolution of products is evident in Figure 5.1, Figure 5.2 and Figure 5.3 (0.5° for TMMM/TRMM RT and 2.5° for NCEP). Due to these differences, comparisons of explained variance across products should be interpreted with care. For example, the relatively coarse resolution of NCEP could mean that sub-grid scale areas, with local strong correlations are being missed.

Table 5.1: Critical values of significance for each basin using their effective sample size having adjusted for autocorrelation (originally 120 observations).

Basin	Autocorrelation coefficient	Effective sample size	Critical rho ($p = 0.05$)
Toktogul	0.74	18	0.47
Andijan	0.75	17	0.49
Kayrakkum	0.74	18	0.47
Nurek	0.80	14	0.54

Figure 5.2: As for Figure 5.1 but with TRMM RT estimated precipitation.

Figure 5.3: As for Figure 5.1 but with NCEP derived precipitation.

Inflows to Toktogul during the summer half-year (April-September mean, n = 9) were significantly related to winter TRMM precipitation averaged across the basin and winter TRMM and NCEP precipitation over the most sensitive sub-basin cell (Table 5.2). Winter NCEP basin average and optimal cell precipitation were significantly related to Andijan summer inflow. TRMM basin average and optimal cell, along with winter TRMM RT optimal cell precipitation, were weakly related to summer Andijan inflows (but were statistically insignificant predictors at p = 0.05). Only TRMM optimal cell was significantly related to summer inflows to Kayrakkum. All sources of winter precipitation tested across the Nurek catchment provided no statistically significant skill for forecasting summer inflows during the fit period.

Table 5.2: Statistical estimates of the intercepts (α) and parameters (β) of simple linear regression models, along with the amount of explained variance (R^2), standard error (SE) of the mean summer (April-September) runoff estimate (m^3s^{-1}) and model significance level (p). All predictors are for the mean winter half-year (October-March), both for the basin average (_Av) and optimal cell identified from correlation surface plots (_Op) above. Values in bold significant at $p = 0.05$.

	Predictor	α (m^3s^{-1})	β	R^2 (%)	SE (m^3s^{-1})	p value
Toktogul	TRMM_Av	**124**	**3.20**	**46**	**118**	**0.04**
	TRMM_Op	**299**	**4.35**	**49**	**115**	**0.04**
	TRMMRT_Av	557	0.32	5	157	0.57
	TRMMRT_Op	621	0.17	2	159	0.69
	NCEPP_Av	227	2.32	26	138	0.16
	NCEPP_Op	**327**	**6.34**	**48**	**116**	**0.04**
Andijan	TRMM_Av	-62	1.13	38	81	0.08
	TRMM_Op	10	1.34	42	78	0.06
	TRMMRT_Av	104	0.21	10	97	0.42
	TRMMRT_Op	-114	0.87	33	84	0.10
	NCEPP_Av	**-206**	**1.57**	**50**	**72**	**0.03**
	NCEPP_Op	**-206**	**1.27**	**56**	**68**	**0.02**
Kayrakkum	TRMM_Av	-132	2.75	33	161	0.10
	TRMM_Op	**138**	**6.05**	**45**	**146**	**0.05**
	TRMMRT_Av	173	0.69	18	178	0.25
	TRMMRT_Op	392	0.25	3	193	0.64
	NCEPP_Av	-165	1.89	16	180	0.29
	NCEPP_Op	-39	0.67	11	185	0.38
Nurek	TRMM_Av	840	1.07	19	95	0.24
	TRMM_Op	882	0.57	20	95	0.23
	TRMMRT_Av	1094	0.05	0	105	0.87
	TRMMRT_Op	1145	-0.06	0	105	0.86
	NCEPP_Av	814	0.72	10	101	0.42
	NCEPP_Op	1361	-0.58	10	100	0.42

The correlation between inflow and all candidate predictors (Table 4.1) for each basin was assessed by systematically varying lead time ($t + 0$ to $t + 4$) and averaging period (1-6 months) (Figure 5.4, Figure Apx 2, Figure Apx 3, Figure Apx 4).

Figure 5.4: Correlation (*r*) of inflows to Toktogul with basin average and optimal cell predictors (leading by 0-4 months) averaged over 1-6 months. Accounting for autocorrelation at *p* = 0.05, *r*ₐ = 0.47.

Each precipitation and temperature product yields two potential predictors, one a basin average (of all cells overlapping the drainage area) and the optimal predictor cell (with greatest explained variance at each lag interval identified from correlation surfaces in Figure 5.1, Figure 5.2, Figure 5.3 and Figure Apx 1). This procedure allowed the averaging period with greatest explained variance in inflows to be identified for each lead time. For example, when forecasting Toktogul discharge at lead time one month (red bars in Figure 5.4) using TRMM basin average, an averaging period of two months yields greatest predictability ($r = 0.83$). Predictors in general transitioned from a positive to negative correlation as lead interval and averaging period increase.

As the lead time and averaging period increase, correlations become negative for most predictors (Figure 5.4). This is perhaps most marked when correlating inflows with NCEP temperature, where strong negative correlations occur for lead times of 3-4 months at averaging periods of 4-6 months. This is likely caused by annual temperature and reservoir inflow regimes being in anti-phase, with the strong correlation driven by the opposing annual regimes rather than a direct influence. This suggests that correlations built using de-trended data may have better reflected the physical mechanisms of runoff generation in the catchment. Interestingly, correlations between NCEP optimal cell precipitation and Toktogul inflows at a three/four-month lead time are stronger than at lead time two months (Figure 5.4). This caused by the location of the optimal NCEP cell precipitation changing between the one/two and three/four month lead time as the area of greatest predictability migrates westward (Figure 5.3).

5.1.2 Model construction and calibration

Once the averaging period with greatest explained variance in inflows was identified for each lead time, stepwise linear regression was performed to achieve the most parsimonious forecasting models. This procedure reduced the risk of over-fitting models to the calibration period. The final sets of predictor variables included in each forecasting model at each lead time are shown in Table 5.3.

Antecedent discharge is only included in models for Toktogul at lead time one month, highlighting the limited autocorrelation in flows for the reservoir inflows studied. Relatively long averaging periods are used for precipitation predictors at lead times 3-4

months for Toktogul. This is attributed to the model attempting to include winter snow accumulation to forecast summer melt driven inflows. Whilst this could work well during summer, it forces winter flows to be predicted from summer precipitation, a mechanism that does not make physical sense. This problem occurs at all reservoirs. No gauge adjusted precipitation product is selected for research models at Andijan. This could be explained by the gauge adjustment procedure being better over the Toktogul basin. Inspection of the network of gauges used by TRMM[26] shows that none are available within TRMM cells overlapping the Andijan basin, compared with 5 for the Toktogul basin during 2000-2010.

Outflows from Toktogul and Andijan were selected for both Kayrakkum models at lead time one month, stressing the importance of water management for forecasting Kayrakkum inflows. It might be expected that model skill at longer lead times is suppressed without this information. The few predictors added to Nurek models may be caused by high level of skill provided by the ZOF. Furthermore, the reliance of operational models at lead times one and three on NCEP precipitation is a cause for concern given the poor relationship between this product and observed precipitation at Sari-Tash.

5.1.3 Model validation

Given the limited amount of inflow data, a k-fold (9 folds used) cross-validation technique was used to benchmark the predictive skill of regression models to the Zero Order Forecast (ZOF) over forecast horizons of $t + 1$ to $t + 4$ months (Q1-Q4). Two model types were tested: (1) 'research grade' models (similar to those built in Dixon and Wilby, 2016) based on historic observations with typical latency 2-3- months; and (2) 'operational' models which only use products available in real time (latency ~2 days). Cross validated model metrics are provided in Table 5.4 for both research and operational models.

Research models were superior to ZOF at lead times of Q1-Q4 for all performance metrics except MARE at Q3 when forecasting Toktogul inflows. This relative improvement in skill can be seen in Figure 5.5, for example during the summers of 2008

---

[26]The GPCC visualizer enables inspection of sites used in TRMM: http://kunden.dwd.de/GPCC/Visualizer

and 2009. For Kayrakkum, research models outperformed the ZOF at all lead times for all performance metrics. When forecasting Andijan inflows, the ZOF was out-performed at all lead times for all metrics except MARE, for which the ZOF was not surpassed at any lead time. This performance is clear during the summer months, but does not reflect the poor skill during winter, where negative inflows were forecast during some years (Figure 5.6). A full set of hydrographs are provided in Appendix A.

Table 5.3: Predictors used in both the operational and research grade multiple regression models for each basin at each lead time (1-4 months) (n = 111). Notations can be found in Table 4.1.

		Q1	Q2	Q3	Q4
Toktogul	Operational	M $Q_{1,1}$ $NCEP\_T\_Op_{1,1}$	M	M $NCEP\_P\_Op_{3,4}$	M $NCEP\_P\_Op_{4,3}$
	Research	M $Q_{1,1}$ $TRMM\_Av_{1,2}$	M $TRMM\_Av_{2,1}$	M $NCEP\_P\_Op_{3,4}$ $TRMM\_Av_{3,1}$	M $NCEP\_P\_Op_{4,3}$ $TRMM\_Av_{4,6}$ $NCEP\_T\_Av_{4,4}$
Andijan	Operational	M $NCEP\_P\_Av_{1,5}$ $TRMM\_RT\_Op_{1,4}$	M $NCEP\_P\_Av_{2,3}$	M $NCEP\_P\_Av_{3,2}$	M $NCEP\_P\_Av_{4,1}$
	Research	M $NCEP\_P\_Av_{1,5}$ $TRMM\_RT\_Op_{1,4}$	M $NCEP\_P\_Av_{2,3}$	M $NCEP\_P\_Av_{3,2}$	M $NCEP\_P\_Av_{4,1}$
Kayrakkum	Operational	M $And_{1,1}$ $Tok_{1,1}$	M $And_{2,1}$	M $NCEP\_P\_Op_{3,4}$ $TRMM\_RT\_Av_{3,6}$	M $TRMM\_RT\_Av_{4,6}$
	Research	M $And_{1,1}$ $Tok_{1,1}$ $TRMM\_Av_{1,5}$	M $TRMM\_Av_{2,4}$	M $NCEP\_P\_Op_{3,4}$ $TRMM\_RT\_Av_{3,6}$	M $TRMM\_RT\_Av_{4,6}$
Nurek	Operational	M $NCEP\_P\_Av_{1,6}$ $NCEP\_P\_Op_{1,6}$	M	M $NCEP\_P\_Op_{3,4}$	M
	Research	M $TRMM\_Av_{1,6}$	M $TRMM\_Av_{2,5}$	M $NCEP\_P\_Op_{3,4}$	M

Table 5.4: Cross validated model metrics for ZOF, Q1, Q2, Q3 and Q4 research and operational models (ZOF used if no predictors added to model during stepwise procedure) (n = 111). Research grade models take as input all products including observation adjusted (e.g. TRMM 3B43, latency of 2-3 months), whereas operational models are composed only of products that would be available to operators in near real-time (latency of ~2 days). Values in bold signify the most promising model for each lead time.

	Metric	Research models					Operational models				
		ZOF	Q1	Q2	Q3	Q4	ZOF	Q1	Q2	Q3	Q4
Toktogul	R^2	0.794	0.871	**0.817**	**0.846**	**0.823**	0.794	**0.874**	ZOF	0.840	0.822
	NSE	0.794	0.871	**0.816**	**0.845**	0.822	0.794	**0.873**	ZOF	0.840	0.822
	AIC	580	558	**575**	568	577	580	**557**	ZOF	568	**574**
	RMSE	152	120	**144**	**132**	**141**	152	**119**	ZOF	134	142
	MARE	20.1	**16.6**	**23.7**	26.6	19.8	20.1	18.8	ZOF	**26.1**	**19.6**
Andijan	R^2	0.555	0.666	0.664	0.654	0.611	0.555	0.666	0.664	0.654	0.611
	NSE	0.550	0.662	0.661	0.651	0.607	0.550	0.662	0.661	0.651	0.607
	AIC	519	507	505	507	514	519	507	505	507	514
	RMSE	88	76	76	78	82	88	76	76	78	82
	MARE	39.8	55.3	54.1	51.6	47.1	39.8	55.3	54.1	51.6	47.1
Kayrakkum	R^2	0.400	**0.584**	0.427	0.432	0.429	0.400	0.563	**0.450**	0.432	0.429
	NSE	0.391	**0.577**	0.414	0.412	0.416	0.391	0.556	**0.443**	0.412	0.416
	AIC	624	**610**	624	626	624	624	611	**621**	626	624
	RMSE	227	**189**	223	223	223	227	194	**217**	223	223
	MARE	33.6	**26.3**	31.8	32.0	31.8	33.6	26.9	**31.6**	32.0	31.8
Nurek	R^2	0.943	**0.943**	**0.943**	0.945	ZOF	0.943	0.939	ZOF	0.945	ZOF
	NSE	0.943	**0.943**	**0.942**	0.945	ZOF	0.943	0.939	ZOF	0.945	ZOF
	AIC	559	**562**	**562**	559	ZOF	559	568	ZOF	559	ZOF
	RMSE	127	**127**	**128**	124	ZOF	127	132	ZOF	124	ZOF
	MARE	14.2	**14.6**	**14.7**	15.4	ZOF	14.2	16.1	ZOF	15.4	ZOF

Figure 5.5: Selected cross validated inflow forecasts for Toktogul, at lead time one (Q1) and three (Q3) months, for operational and research grade models.

Figure 5.6: Selected cross validated inflow forecasts for Andijan, with lead time one (Q1) and two (Q2) months, for the operational and research grade models. Note that both research grade and operational models were built using same predictors.

The challenging standard set by the ZOF for Nurek results in fewer models able to outperform ZOF (Table 5.4). No research model was able to outperform the ZOF at Q1, Q2 and Q4. However, at Q3 the model outperformed the ZOF for all metrics other than MARE.

The performance of operational models (using predictors available in real-time) varies by basin. The best performing models at each lead time for Andijan used only real-time predictors, meaning results are identical to the research models (Table 5.4). A similar situation occurs when forecasting Kayrakkum inflows, with Q3 and Q4 models identical to the research-grade models whilst Q1 and Q2 models outperformed the ZOF for all performance metrics used (note that Toktogul and Andijan outflows were included in Q1 and Q2 models). The operational models were able to outperform ZOF for Toktogul for all metrics (except MARE at Q3) for lead times Q1, Q3 and Q4. At Q2, no additional predictors were added to the model during the stepwise regression process.

Operational models struggled to attain the same performance as the ZOF for Nurek, with no additional predictors added beyond the dummy variable for month at Q2 and Q4. However, the operational model did outperform ZOF at Q3 for all metrics except MARE.

Stratifying models to forecast only summer (Apr-Sep) discharge (using individual months, n = 58), both Toktogul and Andijan outperform the ZOF at all lead times for all metrics used, with the exception of MARE at Q4 (Table 5.5). Similar to the full year investigation, the ZOF of Nurek inflows performs well according to all metrics used, with only modest improvements in skill possible for a few lead times. For Toktogul and Andijan at Q1 and Q2, stratifying models lead to improvements in all metrics with the exception of MARE. An examination of the cross validated hydrographs show stratified models result in generally smaller residuals for both Andijan (Q1) and Toktogul (Q1 and Q2) (Figure 5.7 and Figure 5.8). However, at Q3 stratified models are generally out-performed by their full year equivalents (Figure 5.8). Similar performance metrics are obtained for both full year and stratified models at Q4 (Table 5.5).

Figure 5.7: Cross validated operational inflow forecast residuals for Andijan with lead time one month (Q1). Full year models forecast all 12 months, whereas stratified models forecast only summer inflows (April-September), using ZOF as winter flows.

Table 5.5 Cross validated stratified model metrics for ZOF, Q1, Q2, Q3 and Q4 models for summer months only (April-September) when using full year models (n = 111) versus summer only models (n = 58) (ZOF denotes no predictors added to model during stepwise procedure). Values in bold signify highest performing model for each lead time.

		Full year					Summer only				
	Metric	ZOF	Q1	Q2	Q3	Q4	ZOF	Q1	Q2	Q3	Q4
Toktogul	R2	0.581	0.757	ZOF	**0.720**	0.640	0.581	**0.802**	**0.773**	0.635	**0.649**
	NSE	0.578	0.756	ZOF	**0.720**	0.638	0.578	**0.801**	**0.772**	0.633	**0.643**
	AIC	310	289	ZOF	**291**	**298**	310	**285**	**287**	298	301
	RMSE	214	162	ZOF	**174**	198	214	**147**	**157**	199	**197**
	MARE	23.4	**16.8**	ZOF	**19.8**	**21.7**	23.4	19.7	**19.9**	22.0	27.2
Andijan	R2	0.360	0.602	0.569	**0.544**	0.461	0.360	**0.772**	**0.678**	0.447	**0.484**
	NSE	0.350	0.602	0.569	**0.543**	0.457	0.350	**0.771**	**0.675**	0.439	**0.469**
	AIC	270	260	261	**262**	**267**	270	**246**	**255**	268	268
	RMSE	123	96	100	**103**	112	123	**73**	**87**	114	**111**
	MARE	56.6	46.5	**53.9**	53.9	58.4	56.6	**42.6**	54.7	**49.7**	**49.0**
Nurek	R2	0.812	**0.797**	ZOF	**0.821**	ZOF	0.812	ZOF	ZOF	0.818	ZOF
	NSE	0.812	**0.796**	ZOF	**0.820**	ZOF	0.812	ZOF	ZOF	0.817	ZOF
	AIC	289	**295**	ZOF	290	ZOF	289	ZOF	ZOF	290	ZOF
	RMSE	175	**183**	ZOF	**171**	ZOF	175	ZOF	ZOF	173	ZOF
	MARE	13.6	**14.4**	ZOF	**13.3**	ZOF	13.6	ZOF	ZOF	13.6	ZOF

Figure 5.8: Cross validated operational inflow forecast residuals for Toktogul with lead times one (Q1), two (Q2) and three (Q3) months. Full year models forecast all 12 months, whereas stratified models forecast only summer inflows (April-September), using ZOF as winter flows.

97

5.2   Climate mode based seasonal forecasting

5.2.1   Correlation analysis

Climate modes were selected based on possible relationships to CA precipitation/river flow in previous research, as well as correlating reservoir inflows with global SST's. From this, monthly NAO, Niño 3.4 and DMI indices were selected as predictors of CA reservoir inflows. Monthly indices were correlated with NCEP re-analysis precipitation and temperature across CA for each month during the period 1950-2014, over lead times (index leading) 0 to 3 months.

The NAO was weakly correlated with concurrent NCEP re-analysis precipitation across CA throughout the year over lead times of 1 to 3 months (Figure Apx 20). Only April precipitation across southern Tajikistan showed significant ($p<0.05$) negative correlations with NAO over lead times of 1-3 months. Niño 3.4 showed positive correlations with CA precipitation that were strongest during winter (Figure 5.9). Both January and March showed significant concurrent and lagged positive correlations ($r>0.25$) south-east of the Aral Sea. Significant positive correlations ($r>0.25$) were also present across western Tajikistan and Uzbekistan during October, with stronger correlations ($r\sim0.35$) over Iran and north-western Afghanistan. DMI had only insignificant correlations with precipitation over CA (not shown).

Weak negative correlations ($r\sim0.4$, $p<0.05$) existed between NAO and air temperature across southern CA during January (concurrent only), February (lead 0-1 month) and March (0-2 months) (Figure Apx 21). These correlations were strongest across Tajikistan and northern Afghanistan, but also over parts of Kyrgyzstan. November and December temperature was negatively correlated ($r\sim-0.30$, $p<0.05$) with NAO over north-eastern Iran. Niño 3.4 showed positive correlations ($r\sim0.3$, $p<0.05$) with January and February precipitation over eastern Kyrgyzstan and north-west China at lead times 1-3 months, with insignificant positive correlations across Tajikistan at all lead times (Figure 5.10). Stronger, significant negative correlations ($r\sim-0.4$, $p<0.01$) were present over northern Kyrgyzstan and Kazakhstan during May at lead times 1-3 months. July-September had insignificant negative correlations between Niño 3.4 and re-analysis temperature.

Figure 5.9: Monthly Niño 3.4 correlated with monthly NCEP precipitation rate for the years 1950-2014 at lead times 0-3 months (left column shows target precipitation month, with index month varied to account for lead time). Yellow/dark blue areas are significant at *p*=0.05, orange/pink are significant at *p*=0.01. (Dixon and Wilby, under revision)

Figure 5.10: As in Figure 5.9 but for NCEP temperatures. (Dixon and Wilby, under revision)

Individual winter (October-March) monthly, bi-monthly and half year mean indices were then correlated with mean summer inflows to Toktogul, Andijan and Nurek. The non-parametric Spearman Rank correlation was used to test associations over the periods 1941-1980 and 2001-2010. January and January-February mean NAO were negatively correlated (r~-0.25, p<0.05) with summer inflows to all reservoirs during the period 1941-80 (Table 5.6). December NAO was negatively correlated (r~-0.70, p<0.05) with summer inflows to Toktogul and Andijan during 2001-2010. Winter Niño 3.4 showed strong positive correlations to following summer inflows for all reservoirs over both periods tested. Nino 3.4 was positively correlated (r~0.45, $p < 0.01$) with summer inflows to Toktogul and Andijan for all months except February and March during the period 1941-80. Positive correlations (r ~0.50, $p < 0.01$) were observed for all Niño 3.4 variations with summer inflows at Nurek for the years 1941-1980. Fewer significant correlations were seen during the years 2001-2010, yet correlations remain relatively strong for all months at all reservoirs (r ~0.55, insignificant). Only March DMI showed significant correlations to following summer inflows into Nurek during either period (Table 5.6).

Table 5.6: Spearman rank correlation for individual months and period mean indices versus following April-September inflow anomalies (red= *p* < 0.05, <mark style="background-color:#f8a0a0;color:red">red</mark>= *p* < 0.01). Tok= Toktogul, And= Andijan, Nur= Nurek. (Dixon and Wilby, under revision)

		Oct	Nov	Dec	Jan	Feb	Mar	Oct-Nov	Nov-Dec	Dec-Jan	Jan-Feb	Feb-Mar	Oct-Mar
NAO 1941-80	Tok	0.23	-0.07	-0.10	-0.36	-0.18	-0.07	0.10	-0.11	-0.34	-0.38	-0.14	-0.21
	And	0.12	0.00	-0.10	0.46	0.00	-0.03	0.08	-0.08	-0.41	-0.33	-0.01	-0.16
	Nur	-0.07	0.17	0.06	-0.36	-0.16	0.08	0.11	0.11	-0.24	-0.38	-0.08	-0.15
NAO 2001-10	Tok	-0.08	0.26	-0.69	-0.47	-0.12	0.12	-0.04	-0.60	-0.61	-0.16	-0.02	-0.35
	And	-0.01	0.19	-0.71	-0.50	-0.02	0.07	0.02	-0.62	-0.64	-0.13	-0.01	-0.32
	Nur	-0.44	0.72	-0.26	0.12	-0.45	-0.04	-0.24	-0.10	-0.03	-0.05	-0.26	-0.13
Niño 3.4 1941-80	Tok	0.51	0.52	0.48	0.44	0.43	0.34	0.53	0.48	0.47	0.46	0.43	0.49
	And	0.47	0.52	0.45	0.42	0.39	0.31	0.50	0.46	0.44	0.42	0.39	0.46
	Nur	0.70	0.69	0.67	0.66	0.62	0.52	0.71	0.68	0.69	0.67	0.62	0.68
Niño 3.4 2001-10	Tok	0.50	0.61	0.59	0.54	0.65	0.77	0.56	0.59	0.55	0.58	0.77	0.59
	And	0.42	0.53	0.50	0.45	0.59	0.72	0.48	0.50	0.47	0.50	0.72	0.50
	Nur	0.55	0.60	0.65	0.55	0.54	0.60	0.58	0.65	0.62	0.58	0.60	0.65
DMI 1941-80	Tok	0.24	0.19	-0.01	0.03	-0.09	-0.11	0.24	0.13	0.09	-0.03	-0.07	0.15
	And	0.11	0.05	0.08	0.12	0.00	0.04	0.10	0.09	0.21	0.08	0.04	0.17
	Nur	0.28	0.09	-0.06	-0.16	-0.29	-0.32	0.22	0.03	-0.14	-0.25	-0.28	-0.07
DMI 2001-10	Tok	-0.37	-0.05	0.38	-0.06	0.10	0.05	-0.20	0.24	0.24	0.12	0.10	0.21
	And	-0.39	-0.08	0.31	-0.07	0.04	0.03	-0.22	0.18	0.20	0.07	0.06	0.15
	Nur	-0.36	-0.02	0.33	0.29	-0.27	0.02	-0.26	0.15	0.35	0.07	-0.07	0.02

5.2.2 Composite forecasting

Average summer inflows were pooled conditional on the tercile (i.e. <33%, 33-66%, >66%) of prior winter indices. This enabled evaluation of equal size samples of mean summer inflows conditional on prior climate mode index (negative, neutral or positive). Concurrent observed summer (April-September) temperature was also used to condition summer inflows (cool, average or warm; boundaries provided in Figure 5.11).

The largest difference in summer discharge occurred following opposing Niño 3.4 phases (Figure 5.11). Nurek showed most potential for conditioning by prior indices, with increased (decreased) inflows following El Niño (La Niña) (+15%/-11%) and during warm (cold) summers (+7%/-5%) relative to 1941-1980 baseline respectively.

Figure 5.11: Exceedance probabilities of April-September flow anomalies according to previous October-March index phase (in final column concurrent Apr-Sep temperture used to condition flow) for the period 1941-1980 (0.33 and 0.66 tercile boundaries shown in header).

KS tests showed significant ($p < 0.01$) differences between distributions of inflows to all three reservoirs following El Niño compared with La Niña winters, as well as for inflows to Nurek during cold / warm summers (Table 5.7). The most significant ($p < 0.01$) differences were for summer inflows to Nurek following winter La Niña / El Niño or winter neutral / El Niño conditions. KW tests corroborate these findings with significant differences in summer flows following differing phases of winter Niño 3.4 for all reservoirs ($p < 0.01$); with Nurek showing the greatest difference (Table 5.8).

Table 5.7: Kolmogorov-Smirnov D-statistic based on paired exceedance probability distributions in Figure 5.11 (red**= $p < 0.05$, **red**= $p < 0.01$). (Dixon and Wilby, under revision)**

		Niño 3.4	NAO	DMI	Temperature
Toktogul	Negative - Neutral	0.39	0.23	0.39	0.31
	Neutral - Positive	0.56	0.28	0.28	0.34
	Negative - Positive	0.62	0.33	0.21	0.34
Andijan	Negative - Neutral	0.31	0.31	0.54	0.31
	Neutral - Positive	0.48	0.42	0.35	0.24
	Negative - Positive	0.70	0.27	0.35	0.24
Nurek	Negative - Neutral	0.31	0.23	0.23	0.31
	Neutral - Positive	0.85	0.42	0.36	0.41
	Negative - Positive	0.85	0.40	0.28	0.57

Table 5.8: Kruskal-Wallis H-statistic based on exceedance probability distributions in Figure 5.11 (red**= $p < 0.05$, **red**= $p < 0.01$). (Dixon and Wilby, under revision)**

	Niño 3.4	NAO	DMI	Temp
Toktogul	9.9	2.0	2.6	2.3
Andijan	10.2	3.7	4.7	0.8
Nurek	20.9	2.7	1.0	4.9

Figure 5.12: Summer (April-September) discharge anomalies stratified by preceding October-March Niño 3.4 phase and concurrent air temperature (cold/average/warm) for the period 1941-80. (Dixon and Wilby, under revision)

Focusing on the most promising results from the previous analysis, average summer inflows were evaluated using previous winter Niño 3.4 phase and concurrent temperature anomalies (cold/average/warm) (Figure 5.12). Nurek had higher (lower) than average inflows following El Niño (La Niña) winters, and higher (lower) in warm (cool) summers. Both Toktogul and Andijan exhibit the same pattern, albeit weaker, following each Niño 3.4 phase. However, these reservoirs experienced lower (higher) summer inflows under concurrent warm (cool) conditions.

The stationarity of the Niño 3.4-Nurek inflow relationship was examined using a reconstructed discharge record created for Komsomolabad (based on a seasonal regression with flows gauged at the upstream station of Garm, as described in section 4.3.2). Nurek summer inflows were again conditioned on prior winter Niño 3.4 phase for the period 1981-2016 (years 1993-1999 missing) to assess stationarity with respect to the 1941-1980 data. As before, the period 1981-2016 has increased (decreased) inflows following El Niño (La Niña) (+4%/-7%) (Figure 5.13). However, differences are not statistically significant according to the Kolmogorov-Smirnov and Kruskal-Wallis tests (Table 5.9 and Table 5.10). Using the full discharge series (1941-2016 with 1993-1999 missing), both tests showed significant differences ($p<0.01$) in Nurek inflows following La Niña/El Niño.

Figure 5.13: Exceedance probabilities of April-September flow anomalies (m³s⁻¹) into Nurek according to previous October-March Niño 3.4 phase for differing time periods (tercile based on period 1941-80). (Dixon and Wilby, under revision)

Table 5.9: Kolmogorov-Smirnov D-statistic based on paired exceedance probability distributions for Nurek shown in Figure 5.13 (red= *p* < 0.05, red= *p* < 0.01). (Dixon and Wilby, under revision)

	1941-1980	1981-2016 (1993-1999 missing)	1941-2016 (1993-1999 missing)
Negative - Neutral	0.39	0.29	0.30
Neutral - Positive	0.71	0.42	0.57
Negative - Positive	0.79	0.42	0.61

Table 5.10: Kruskal-Wallis H-statistic based on exceedance probability distributions for Nurek shown in Figure 5.13 (red= *p* < 0.05, red= *p* < 0.01). (Dixon and Wilby, under revision)

	1941-1980	1981-2016 (1993-1999 missing)	1941-2016 (1993-1999 missing)
H-Statistic	19.4	4.9	22.8

As these forecasts used October-March Niño 3.4, the earliest a forecast could be issued would be April after March Niño 3.4 information becomes available. After consultation with the operators of Nurek, it became clear that a forecast capable of being issued in January as well as April was desired. Therefore, the method was developed to use November-December Niño 3.4 to forecast following summer (April-September) inflows to Nurek, allowing an 'early outlook' forecast to be issued in January. Similar differences were seen between summer inflows conditioned on prior November-December and October-March Niño 3.4 (Table 5.11 and Table 5.12). According to KS tests, differences in summer inflows following contrasting phases of Niño 3.4 were greater during 1941-80 when conditioned on November-December Niño 3.4, but less for both 1981-2016 and 1941-2016 periods. Both forecast issue times provide the same level of skill (according to the HHP) when tested on the period 1975-2016 (1993-1999 missing), but November-December outperforms October-March when tested on the period 1941-1974 (0.59 compared to 0.53 respectively) (Table 5.13).

An operational forecast can also be provided for decision makers using a likelihood lookup table (Table 5.14). Using this format, following La Niña conditions during October-March, summer inflow anomalies will most likely range between -199 and -100 m^{3}s^{-1} (with a likelihood of 33%). Following neutral conditions, summer inflows will most

likely fall between 0 to -99 m^3s^{-1} (with a likelihood of 53%) and between +1 to +100 m^3s^{-1} following El Niño (with 35% likelihood). Furthermore, the likelihood of more extreme conditions can be forecast using Table 5.14. For example, following El Niño conditions during October-March there would be 12% likelihood of inflow anomalies exceeding 200 m^3s^{-1} (bottom right of Table 5.14). Using both October-March and November-December Niño 3.4 gives similar results, with the greatest likelihood of flows occurring within the same range following each phase (with the exception of equal likelihood of both -99 to -100 m^3s^{-1} and -99 to 0 m^3s^{-1} following La Niña conditions during November-December).

Table 5.11: As for Table 5.9 but using November-December Niño 3.4 to condition April-September Nurek inflow anomalies. (Dixon and Wilby, under revision)

	1941-1980	1981-2016 (1993-1999 missing)	1941-2016 (1993-1999 missing)
Negative - Neutral	0.39	0.17	0.27
Neutral - Positive	0.70	0.37	0.55
Negative - Positive	0.85	0.39	0.59

Table 5.12: As for Table 5.10 but using November-December Niño 3.4 to condition April-september nurek inflow anomalies. (Dixon and Wilby, under revision)

	1941-1980	1981-2016 (1993-1999 missing)	1941-2016 (1993-1999 missing)
H-Statistic	20.9	3.4	20.9

Table 5.13: Heidke Hit Proportion skill scores for mean April-September Nurek inflow anomalies conditioned on preceding November-December and October-March Niño 3.4. (Dixon and Wilby, under revision)

Calibration period	Validation period	November-December Niño 3.4	October-March Niño 3.4
1941-1974	1975-2016 (1993-1999 missing)	0.51	0.51
1975-2016 (1993-1999 missing)	1941-1974	0.59	0.53

Table 5.14: Likelihood (%) of mean April-September Nurek inflow anomalies (m^3s^{-1}) depending on prior October-March and November-December Niño 3.4 phase for the period 1941-2016 (1993-1999 missing) (all Niño 3.4 terciles are based on the period 1941-1980). (Dixon and Wilby, under revision)

Inflow anomaly (m^3s^{-1})	November-December Niño 3.4 index			October-March Niño 3.4 index		
	La Niña	Neutral	El Niño	La Niña	Neutral	El Niño
≤ -200	9%	5%	0%	13%	0%	0%
-199 to -100	35%	16%	4%	33%	16%	4%
-99 to 0	35%	42%	26%	29%	53%	23%
1 to 100	22%	21%	33%	25%	16%	35%
101 to 200	0%	11%	26%	0%	11%	27%
> 200	0%	5%	11%	0%	5%	12%

Whilst significant differences were present between inflows during cool/warm summers during the period 1941-1980, this relationship does not hold for the periods 1981-2016 or 1941-2016 (not shown). Furthermore, the effect of concurrent temperature (even during 1941-1980) was sensitive to the choice of station used. Figure 5.14 shows the significant differences ($p < 0.01$) in summer inflows to Nurek following warm/cool summers measured at Sari-Tash (3100m) do not hold when using the Lyakhsh (2000m) meteorological station (utilising a blended temperature record of observed and predicted temperature obtained by regressing against Garm as described in section 4.3.2). The sensitivity of the Nurek-Niño 3.4 relationship was also assessed in terms of the period used to derive Niño 3.4 tercile boundaries. When terciles were based on the period 1982-2010, stronger/weaker Niño 3.4 index values were required to be classed as El Niño/La Niña than when using 1941-1980. This caused a greater number of seasons to be classed as neutral resulting in greater discrimination between summer flows when terciles were based on 1982-2010 (Figure 5.15). No matter which period was used to base tercile boundaries, differences in summer inflows following neutral-positive and negative-positive Niño 3.4 phases were significant ($p < 0.01$) according to KS tests.

Figure 5.14: Exceedance probabilities of April-September flow anomalies into Nurek according to concurrent temperature for the period 1941-1980 (flow anamolies and terciles relative to 1941-1980 baseline). (Dixon and Wilby, under revision)

Figure 5.15: Exceedance probabilities of April-September flow anomalies into Nurek (using blended discharge record) according to November-December Niño 3.4 phase for the period 1941-2016 (1993-1999 missing) using terciles based on period 1941-1980 and 1982-2010 (flow anamolies relative to 1941-1980 baseline). (Dixon and Wilby, under revision)

5.3 Transferability of methodologies

The transferability of both approaches was assessed by implementation to the M'Jara catchment in Morocco. This allows an assessment of the capabilities of each approach outside the region of Central Asia.

5.3.1 Statistical forecasting model

The amount of explained variance in monthly inflows explained by overlaying cells of NCEP precipitation and temperature at lead times 0-3 months are shown in Figure 5.16. NCEP derived precipitation was significantly correlated with M'Jara flow at lead times 0-2 months. The area of greatest predictability was across the west and north-west of the basin, with $r > 0.6$ ($p < 0.01$) at lead time one month. NCEP temperature showed strong positive correlations with M'Jara flow at all lead times examined (Figure 5.16). The region of strongest correlation migrated from north-west to south-east as lead time increases. Predictors were systematically varied by lead time (t + 0 to t + 3) and averaging period (1-6 months) to identify averaging periods which provided greatest predictive power at each lead time (Figure 5.17). For lead times one and two months the strongest correlation was achieved with an averaging period of only 1 month for all predictors. At a lead time of three months, strongest correlations occurred with an averaging period of 6 months.

Figure 5.16: Monthly NCEP precipitation and temperature correlated with observed monthly discharge at M'Jara (lagged by 0-3 months) for the period 1952-89 (n = 456). At $p = 0.05$, $r_{crit} = 0.20$.

Figure 5.17: Spearman's rank correlation of monthly discharge at M'Jara (lagged by 0-3 months) with NCEP predictors averaged over 1-6 months for the period 1952-1989. At *p* = 0.05, accounting for autocorrelation r_{crit} = 0.20.

Using the averaging period with greatest explained variance of M'Jara flow for each lead time, stepwise linear regression was performed to achieve the most parsimonious forecasting models. The final sets of predictor variables included in each forecasting model at each lead time are shown in Table 5.15. Only at a lead time of one month (Q1) were predictor variables added to the model beyond the dummy variable for month. Figure 5.18 shows the marginal improvement in forecast skill achieved by the model compared with the ZOF. The model was able to slightly reduce large errors, for example during years 1969, 1970 and 1977. Both the model and ZOF were not able to capture the large flow during January 1970, then over predicted flows the following month. This mismatch of peaks by the model was seen several times throughout the record, for example in December 1968 and December 1963. These timing errors during large

events are reflected in the low R^2 and NSE scores, suggesting that high flows are poorly forecast by both the ZOF and the model.

Table 5.15: Predictors used in multiple regression models at lead times 1-3 months. Predictor notations can be found in Table 4.1.

Lead time	Predictors
ZOF	M
Q1	M, Flow lag$_{1,1}$, NCEP_P_Op$_{1,1}$
Q2	M
Q3	M

Figure 5.18: Cross validated flow forecast residuals at M'Jara for lead time one month (Q1).

Table 5.16: Cross validated model metrics for ZOF, Q1, Q2 and Q3 models (ZOF used if no predictors added to model during stepwise procedure) (n = 456).

Metric	Lead time			
	ZOF	Q1	Q2	Q3
R2	0.22	0.30	ZOF	ZOF
NSE	0.22	0.30	ZOF	ZOF
AIC	2282	2261	ZOF	ZOF
RMSE	142	134	ZOF	ZOF
MARE	8.1	6.6	ZOF	ZOF

5.3.2 Climate mode based approach

Correlations between NAO/Niño 3.4 and precipitation across northern Morocco were investigated at lead times 0-3 months (index leading). Figure 5.19 shows NAO is significantly negatively correlated with concurrent precipitation across north western Morocco during winter. Correlation strength reduced rapidly with distance inland, typically from $r > -0.6$ at the coast to $r < -0.3$ close to the border with Algeria. The strong negative correlation became weak as soon as any lead time is introduced. Niño 3.4 was significantly positively correlated to precipitation during July-October at all lead times (Figure Apx 22). Significant negative correlations occurred between Niño 3.4 and precipitation during March-April, but correlation strength reduces quickly from the coast inland. Only November NAO was significantly negatively correlated ($p < 0.05$) with December-January mean flow at M'Jara (Table 5.17). Both December and January individual ($p < 0.05$) and bi-monthly mean ($p < 0.01$) NAO were significantly negatively correlated with February-March mean flow at M'Jara. Niño 3.4 was not significantly correlated with either December-January or February-March mean flow at any lead times investigated (not shown).

The two target flow periods (December-January and February-April mean) were conditioned on prior NAO phase. Only NAO was used for conditioning due to Niño 3.4 being weakly correlated with flow at M'Jara. Lead times for conditioning were shorter than for CA, using the bi-monthly mean for the two months prior to the target period. This reflected the quickly diminishing relationship between NAO and precipitation/flow, as depicted in Figure 5.19 and Table 5.17. Whilst negative flow anomalies during both target flow periods occurred following any NAO phase, large positive anomalies were more likely to occur following negative phases (Figure 5.20). Only two small positive February-April flow anomalies occurred following positive NAO during December-January. No significant differences were found between conditioned flows using either KS or KW tests for December-January. Significant differences were present between February-April mean flows at M'Jara conditioned on prior December-January NAO according to the KW test ($p < 0.05$) (not shown). KS tests showed significant differences between flows following negative-neutral ($p < 0.05$) and negative-positive ($p < 0.01$) NAO (Table 5.18). Table 5.19 shows there is a 38% likelihood of February-April mean flow anomalies exceeding +200 m^3s^{-1} following negative NAO, compared with 0% following neutral or positive NAO.

Figure 5.19: Monthly NAO correlated with monthly NCEP precipitation rate for the period 1952-1989 at lead times 0-3 months (left column shows target precipitation month, with index month varied to account for lead time). Purple/Yellow areas and greater are significant at p=0.05.

Table 5.17: Spearman rank correlation for individual and bi-monthly NAO versus following December-January and February-April flow anomalies at M'Jara for the period 1952-1989 (red= $p < 0.05$, red= $p < 0.01$).

Target flow months	Index month(s)				
	Sep	Oct	Nov	Sep-Oct	Oct-Nov
December-January	-0.26	0.14	-0.35	0.07	-0.07
	Nov	Dec	Jan	Nov-Dec	Dec-Jan
February-April	-0.07	-0.41	-0.36	-0.30	-0.46

Figure 5.20: Exceedance probabilities of flow anomalies (m^3s^{-1}) at M'Jara according to previous NAO phase for the period 1952-1989.

Table 5.18: Kolmogorov-Smirnov D-statistic based on paired exceedance probability distributions shown in Figure 5.20 (red= $p < 0.05$, red= $p < 0.01$).

	December-January	February-April
Negative - Neutral	0.42	0.54
Neutral - Positive	0.35	0.43
Negative - Positive	0.31	0.62

Table 5.19: Likelihood (%) of mean February-April flow anomalies at M'Jara (m^3s^{-1}) depending on prior December-January NAO phase for the period 1952-1989.

Inflow anomaly (m^3s^{-1})	December-January NAO phase		
	Negative	Neutral	Positive
≤ -100	8%	33%	38%
-99 to 0	31%	17%	46%
1 to 100	8%	42%	15%
101 to 200	15%	8%	0%
> 200	38%	0%	0%

5.4   Summary

The overall aim of this research is to develop data streams and model structures that facilitate reservoir inflow forecasting in data sparse regions using public domain predictors. This chapter has presented both approaches used to achieve this aim, including predictor selection, followed by the construction, calibration and validation of statistical models. Both approaches have also been implemented for a basin in Morocco to allow assessment of approach transferability.

TRMM precipitation estimates consistently explain more variance in CA reservoir inflows than TRMM RT. NCEP precipitation correlates with flow at strengths comparable to TRMM across lead times 0-3 months. However, the coarse resolution of NCEP precipitation (2.5°) could mean that fine scale variations in correlation strength may be lost. The effect of introducing an averaging period of 0-6 months was evaluated for basin average and optimal cell precipitation and temperature predictors. These procedures identified the locations and averaging periods of predictors yielding greatest explained variance in inflows to each reservoir at lead times 1-4 months. Models were then built using all available significant predictors (research grade) and significant predictors available in near real-time (operational). Overall, both research grade and operational models were able to outperform the ZOF for Toktogul, Andijan and Kayrakkum for all lead times using most skill metrics. Stratifying models to forecast only summer inflows resulted in greater skill at lead times 1, 2 and 4 months for Toktogul and Andijan.

Niño 3.4 was positively correlated with winter precipitation across Tajikistan and Kyrgyzstan. Significant correlations ($p < 0.05$) occurred during October and March at lead times 0-3 months for the period 1950-2014. Summer inflows to Toktogul, Andijan ($p < 0.05$) and Nurek ($p < 0.01$) were significantly correlated with all prior winter months Niño 3.4 during the period 1941-1980. Furthermore, significant differences ($p < 0.01$) were present between summer inflows to all three reservoirs conditioned on prior October-March Niño 3.4, as well as Nurek inflows conditioned on concurrent temperature for the same period. However, no significant differences were present amongst inflows to Nurek when replicating the method using the period 1981-2016. Using the November-December mean Niño 3.4 allows a forecast to be issued in January, and has similar discriminatory power compared to using October-March mean Niño 3.4. Likelihood (%) tables provide a method of displaying the uncertainty in forecasts to operators/decision makers.

The transferability of approaches was then assessed, first using the statistical approach. NCEP precipitation and temperature were significantly correlated to M'Jara flow at lead times 0-2 months. Models were again built using optimal cell and basin average predictors at averaging periods yielding greatest predictability at lead times 0-3 months. Only at a lead time of one month were any variables included in the regression models in addition to the dummy variable for month. Evaluation of model residuals showed a number of very large errors, generally under-estimating large flows followed by over-estimating the following months flow.

NAO was significantly correlated with concurrent northern Morocco precipitation from September through to May. Correlation strength reduces distinctly as soon as one-month lead time is introduced, as well as with distance inland. Two forecast target periods were used, December-January and February-April means. December-January NAO was significantly correlated ($p < 0.01$) with February-April flow at M'Jara. Significant differences ($p < 0.01$) were also present between February-April mean flow following negative/positive phases of December-January NAO. The next chapter discusses the results presented here, allowing an evaluation of findings in relation to prior literature as well as offering physical interpretations for the statistical relationships identified.

6 Discussion

The previous chapter presented the results of thesis, produced using the methodologies presented in Chapter 4. Two forecasting approaches were implemented, a statistical regression-based method and climate mode-based forecast. Both approaches were then assessed for their transferability via implementation for the M'Jara catchment in Morocco. The headline findings from the chapter include that inflows to Toktogul and Andijan reservoirs were better forecast than Kayrakkum and Nurek inflows using the statistical forecasting approach. Improvements in skill were achieved for Toktogul and Andijan by stratifying models to forecast inflows during only summer months (April-September). Significant differences ($p < 0.01$) were present in mean summer inflows to Toktogul, Andijan and Nurek reservoirs following opposing phases of winter Niño 3.4 during the years 1941-1980. Winter Niño 3.4 showed greatest discriminatory power for summer inflows to Nurek reservoir. Significant differences did not hold however for the period 1981-2016 (1993-1999 missing). Using the November-December mean Niño 3.4 allows a forecast of Nurek inflows to be issued in January, whilst showing greater skill according to HHP in forecasting summer inflows than October-March mean Niño 3.4. The statistical forecasting approach had little skill (NSE of 0.30 at one-month lead time) at forecasting flows at M'Jara. The climate mode-based approach showed greater potential for transferability to Morocco, with statistically significant differences between February-April flows following opposing phases of December-January NAO.

This chapter will contextualise the findings, drawing comparisons within each forecasting approach regarding relative performance for each reservoir and providing physical insights to explain statistical findings. Results will also be discussed in relation to prior literature, with a focus on any contrasting conclusions. The statistical forecasting approach will be discussed first, before the climate mode-based approach. A physical mechanism is also presented that could explain how ENSO affects CA streamflow. The transferability of both modelling approaches is then assessed, using the M'Jara catchment in Morocco as a test basin. Furthermore, the potential for operationalising a seasonal reservoir inflow forecast for Nurek reservoir in CA will be evaluated. The strengths and weaknesses of both forecasting approaches to CA reservoir inflows as well as their transferability to the M'Jara basin in Morocco are evaluated. Finally, the

operational potential of the climate mode-based approach is discussed, providing scope for the forecasting of other predictors from ENSO.

6.1   Statistical forecasting models

The statistical forecasting approach used regression-based models to relate precipitation and temperature products to reservoir inflows. First, correlations between precipitation/temperature products and inflows were evaluated, enabling the identification of cells/areas of greatest potential predictability of inflows at lead times 0-4 months. Multiple linear regression models were then built using a stepwise procedure, with the basin average and highest performing (optimal) cell at each lead time available for selection. Two types of model were created: 'research' grade and an 'operational' version (Table 4.2). Research grade models included all available predictors, whereas operational versions used only those predictors that would be available in (near) real-time. The operational and research grade model skill at forecasting Toktogul inflows are first evaluated in relation to the findings of Dixon and Wilby (2016). The relative performance of operational and research grade statistical models at each reservoir will then be discussed. This allows an interpretation of statistical findings using the physical understanding provided in section 3.1.1 to extract the drivers of predictability. Operational models were then stratified to forecast only summer inflows, the period of peak natural inflows. Finally, the compliance of models with the assumptions of regression is assessed.

6.1.1   Predictor selection

Dixon and Wilby (2016) found a downstream migration of the strongest correlations with increasing lead time from the high Tien Shan Mountains in the East to lower elevation areas in the west of the Toktogul basin, Kyrgyzstan. All precipitation products show this downstream migration for Toktogul inflows, as does TRMM in the Andijan basin (Figure 5.1, Figure 5.2 and Figure 5.3). This behaviour is attributed to variations in the annual precipitation regime across the catchment (Figure 6.1). The high Tien Shan Mountains in the east of the catchment receives maximum precipitation in July compared with April/May in the west (according to TRMM). Nurek shows negative correlations transition to positive as the lead time increases from 0-4 months when using TRMM and

NCEP precipitation. This is likely a result of phase differences in the precipitation and flow regimes of the basin, with precipitation peaking during March/April/May and flow in July. Kayrakkum has a clear spatial transition from positive to negative correlations in the lowlands/headwaters of the basin concurrently. This is caused by considerable regulation of flow above Kayrakkum, with the transition of correlations occurring at the site of Toktogul reservoir (Savoskul et al., 2003; Lutz et al., 2013).

Figure 6.1: Precipitation regime differences in the east and west of the Toktogul basin according to TRMM precipitation for the years 1998-2014.

Schär et al. (2004) assumed a linear relationship between December-April accumulated precipitation (re-analysis and observed) and May-September river flows in both the Syr and Amu Darya rivers. Significant predictions of summer inflow to Toktogul were achieved using both TRMM basin average and optimal cell winter precipitation as well as NCEP optimal cell precipitation (Table 5.2). Summer inflows to Andijan are also significantly related to previous winter NCEP basin average and optimal cell precipitation. Whilst significant, the present correlations were weaker than those achieved by Schär et al. (2004) for the Syr Darya. This could be due to differences in precipitation data, period of record (Schär used 1979-1993) and/or that different flow gauging stations were used with only the headwater regions of the Syr Darya used in this study. Schär et al. (2004) found that winter precipitation had a stronger relationship

to summer flow in the Syr Darya compared with the Amu Darya, aligning well with the results obtained here. Possible explanations are the relatively poor accuracy of precipitation products over the Nurek basin found by Schär et al. (2004) and shown in Table 3.4, as well as the greater proportion of glacier derived inflows to Nurek compared with the Syr Darya (Savoskul and Smakhtin, 2013 and Figure 3.8).

6.1.2 Operational potential of the Dixon and Wilby (2016) methodology

Use of a stepwise procedure with statistically derived input and output values differs from that applied by Dixon and Wilby (2016). Their predictors were manually entered into the model until no further improvement in model skill was achieved via further predictors (by reliance on expert opinion). The F-in and F-out statistics used in this thesis result in the selection of parsimonious models. Parsimony is valued because of the reduced likelihood of 'over-fitting' the model to the training set, resulting in better performance when used as a forecasting tool (Wilks, 2011). Cross validated skill metrics achieved here using research grade models are comparable to those in Dixon and Wilby (2016) for a lead time of one month. At lead time two months, Dixon and Wilby's (2016) model outperforms the model presented here, likely due to antecedent discharge being an additional predictor. At lead time three months, the research model presented here outperforms Dixon and Wilby's (2016) using all skill metrics. It seems the addition of NCEP optimal cell precipitation ($r = 0.83$) provides this extra skill, allowing more skilled forecasts at increased lead times. Temperature was only included at the longest lead time of four months. This supports the assumption that the monthly dummy variable captures most of the seasonal snow and ice melt regime, with limited extra predictability of inter-annual variability in flows by temperature (Dixon and Wilby, 2016).

The use of TRMM precipitation (latency 2-3 months) by Dixon and Wilby (2016) limits the operational potential of their seasonal forecasting models. In this thesis, both research grade and operationally viable models were produced allowing comparison. At lead times one, three and four months both research grade and operational models perform comparably for Toktogul with only minor differences in skill metrics. Both models perform poorly at a two-month lead time compared with Dixon and Wilby (2016), but the operational model cannot outperform the ZOF unlike the research grade model.

The limited predictability of Toktogul inflows at a lead time of two months seems to occur due to the location of the optimal NCEP cell precipitation changing between the one/two and three/four month lead time as the area of greatest predictability migrates westward (Figure 5.3). This change of optimal cell location results in correlations between NCEP optimal cell and Toktogul inflows at a three/four-month lead time being stronger than at lead time two months (Figure 5.4). This, as well as NCEP optimal cell correlations remaining strong as averaging period increases for lead times three/four months, results in greater model skill at lead times of three/four months compared to two months. However, were the system operationalised, all predictors available at the time of forecast issue (i.e. those from three and four-month lead times as well as two months) would also be eligible for inclusion in the model. Therefore, at a two-month lead model skill would be at least as good as models produced at lead times three and four months. Overall, the use of NCEP precipitation and temperature products compared with TRMM precipitation estimates does not result in a significant reduction in model skill. As such, the method of Dixon and Wilby (2016) shows potential for operational implementation.

Model performance can be examined by visual inspection of the forecasted hydrographs benchmarked against both observed inflows and the long-term average (ZOF). At a one-month lead time, both the research and the operational models perform similarly for Toktogul (Figure 5.5). Whilst both model types outperform the ZOF over the record period, in some years the ZOF is more closely aligned with observed peak summer inflows, for example during summer 2009. This could be caused by the model's reliance on antecedent discharge rather than accounting for prior winter precipitation accumulation. Often model improvement over ZOF is marginal with large errors in forecasted annual peak inflow still occurring, as in 2007 and 2010 for example. Again, not fully accounting for winter snow accumulation could be causing this. At a three-month lead time, the poor performance of modelled inflows to Toktogul during winter is clear (Figure 5.5). During the winter months, ZOF consistently outperforms both research grade and operational models. The latter models cannot discriminate between precipitation falling as snow (and therefore being stored) and rain (which is available immediately for flow generation). Hence, inclusion of TRMM and NCEP precipitation actually reduces model performance for low flows during winter. For unregulated flows, it could be argued that performance during the winter is not as important as during

summer because >75% of annual average inflow occurs during April-September for Toktogul, Andijan and Nurek. A case could be made for forecasting only summer inflows, presuming winter low flows to be equal to the long-term average. Such a stratified model may allow greater summer predictability to be drawn from the relationship because it is not being weakened by the inclusion of ambiguous winter precipitation.

6.1.3 Transferability within Central Asia

Dixon and Wilby (2016) investigated the potential for forecasting inflows to Toktogul reservoir. However, there are other large reservoirs in Central Asia for which seasonal forecasts could be beneficial. Therefore, the approach was extended to three other reservoirs in the region: Andijan, Kayrakkum and Nurek. As before, TRMM estimated precipitation and NCEP precipitation and temperature products were available as inputs into the stepwise regression procedure. Each reservoir differs in terms of the timing and volume of inflows, as well as the relative contributions of ice and snowmelt. These differences were used to deepen understanding of model performance.

The operational and research grade models had the same predictors for Andijan (Table 5.3). Unlike Toktogul, gauge adjusted TRMM was not included at any lead time for forecasting inflows to Andijan. This could be explained by the gauge adjustment procedure being better over the Toktogul basin. Inspection of the network of gauges used by TRMM[27] shows that none are available within TRMM cells overlapping the Andijan basin, compared with 5 for the Toktogul basin during 2000-2010. Both models outperform the ZOF at all lead times for all metrics except MARE for Andijan (Table 5.4). This can be seen in the forecasted hydrographs, with modelled inflows being generally closer to observed values than the ZOF, clearly seen during peak flows in summer 2003 and 2008 (Figure 5.6). However, the low summer peak during summer 2007 was forecast poorly at all lead times. It is possible that the dummy variable for month is dominating the precipitation input, reducing the variance of the forecast compared with observed inflows. This could also help to explain the under-prediction of summer 2010 inflows.

---

[27]The GPCC visualizer enables inspection of sites used in TRMM: http://kunden.dwd.de/GPCC/Visualizer

An artefact of the regression model is that negative inflows are forecast into Andijan during winter at all lead times during the period 2001-2010. Similarly to Toktogul, it seems that the long term average actually performs better than the model for the low flow periods. The model seems to be tailored to perform better during the summer months, as the precipitation input uses a long averaging period to best capture winter snowpack accumulation (Table 5.3). A side-effect of this is that winter flows are predicted from summer precipitation, a mechanism that does not make physical sense. This adds weight to the argument that the statistical model should not be used to forecast year-round flows. These negative values have been kept in the plotted hydrographs for transparency, even though they are clearly not possible. It should be noted, however, that these negative flows would have to be addressed if the model was used for operational forecasting. Possibilities include the use of ZOF whenever a negative inflow is forecast, replacing negative values with zero or fitting the model to forecast only winter flows.

Both operational and research grade models also outperform the ZOF when forecasting Kayrakkum inflows using all metrics at all lead times, except AIC after a one-month lead (Table 5.4). Of note is the large reduction in performance of both models when lead times extend beyond one month. Whilst Andijan outflows were still included in the operational model at lead time two months, it is likely that most release from upstream reservoirs has either already entered Kayrakkum, has been stored in subsequent reservoirs or has been used for irrigation by this point. The reliance on Toktogul and Andijan outflows for prediction skill at Kayrakkum is shown during the subdued inflows of 2008/9 and the variable inflows during 2010, which are not well depicted by either model after a lead time of one month (not shown). However, no model forecasts the large inflow during June 2003. From closer inspection of outflows from Toktogul and Andijan there appears to be no anomalous outflows. Therefore, it is surmised that the inflow is due to erroneous data or a large release from one of the reservoirs below Toktogul on the 'Naryn cascade' (e.g. Uch-Kurgansk), but no information on this is available. These results suggest that the management of water above Kayrakkum is the main driver of this reservoir's inflow regime, rather than flows from unregulated sub-catchments. Therefore, accurate forecasts of Kayrakkum inflows requires detailed information on upstream water management (e.g. irrigation schedules and reservoir operation rules), rather than precipitation and temperature alone.

The ZOF proves to be a strong benchmark for forecasting inflows to Nurek reservoir, with a Nash-Sutcliffe efficiency of 0.93 (Table 5.4). Only at lead time three months can models outperform the ZOF, even though additional predictors were added to the model at lead times one and two months. The additional skill at lead time three months is provided by the optimal NCEP precipitation cell averaged over four months, which is located in the eastern headwaters of the basin. This lead time, averaging period and location combination likely allows winter snowpack accumulation to be used to forecast summer inflows. Overall, the ZOF performs to such a high standard that inclusion of other predictors in the models results in only minor variations in modelled inflow. The weaker performance of models for Nurek (in southern CA/Pamirs) compared with Toktogul and Andijan (northern CA/Tien Shan) is consistent with the findings of Schär et al. (2004). The comparatively weak skill of Amu Darya seasonal flow forecasts compared with those for the Syr Darya was explained by less accurate precipitation products and lower quality natural discharge estimates by Schär et al. (2004). It is plausible that less accurate precipitation products contribute to the reduced skill presented here as both TRMM RT and NCEP monthly precipitation are not significantly correlated with observed precipitation at Sari-Tash (Table 3.4).

6.1.4 Stratified models

Due to the clear differences in the key drivers of flow between summer and winter, models were stratified to forecast inflows during summer only (April-September), accepting the long-term monthly mean as winter inflows. This aimed to both reduce the problems of forecasting winter flows, as well as improve forecast skill during the important summer months. Due to upstream regulation rather than natural processes dominating the inflow regime to Kayrakkum, no further model development was undertaken for this reservoir.

For both Toktogul and Andijan reservoirs stratified models outperform full year models at lead times one, two and four months for all metrics except MARE (Table 5.5). As MARE is sensitive to errors occurring at low discharges, it is not of significant importance for forecasting summer peak inflows. Conversely, the other four metrics are sensitive to errors occurring at higher magnitude discharges (see section 4.2.3),

arguably more important for CA as a large proportion of flow occurs during the summer peak. At a lead time of three months, summer only models for Toktogul and Andijan include NCEP precipitation lagged by three months with no averaging period. Full year models on the other hand include the same NCEP precipitation but averaged over four (Toktogul) and two (Andijan) months. It is known that winter precipitation is a skilful predictor of summer inflows in the Syr Darya (Schär et al., 2004; Table 5.2). This extra information regarding winter precipitation (and by inference snowpack accumulation) is, therefore, the likely reason that full year models outperform summer only models at lead time three months. Summer only models at lead times one and two months for both Toktogul and Andijan utilise a relatively long averaging period for their precipitation input (between 3-6 months), providing skill from the previous winter's accumulated precipitation. The ZOF for both Toktogul and Andijan is weaker using only summer months, as would be expected and summer only models outperform the ZOF at all lead times for all metrics except MARE.

The large residuals in forecasted inflows to Andijan during the summers of 2007 and 2010 are reduced by around half when using stratified compared with full year models (Figure 5.7). However, negative inflows are still forecast by stratified models at lead times one, two and four months for September. As inflows to Andijan are snow- rather than ice-melt dominated (Savoskul and Smakhtin, 2013), summer inflows are moisture limited rather than energy (temperature) limited. It is likely that with higher mean air temperatures, snow melt occurs earlier in the year (Sorg et al., 2012; Siegfried et al., 2012). In such circumstances, stratified models could be built for a shorter or earlier period of the year, possibly leading to improvements in model skill for the important summer peak. At lead times one and two months, stratified models generally forecast the summer inflows to Toktogul more accurately (smaller residuals) than full year models (Figure 5.8). However, as shown by the skill metrics, full year models outperform stratified models for a lead time of three months. Stratifying models did not improve forecasts of Nurek inflows, with additional predictors (on top of the monthly dummy variable) only included at lead time three months. The ZOF continued to be a strong benchmark when forecasting only summer inflows with a NSE of 0.81. It is possible that either the poor accuracy of precipitation products over the region (Schär et al., 2004) (Table 3.4) and/or the greater relative importance of glacier melt (Savoskul

and Smakhtin, 2013) reduce precipitation products skill for forecasting summer inflows at this site.

The significant positive correlations between mean winter precipitation (TRMM and NCEP derived) and mean summer runoff found for both Toktogul and Andijan basins are of comparable strength to similar forecasting attempts in the region (Schär et al., 2004; Archer and Fowler, 2008; Barlow and Tippett, 2008). Previous studies have generally forecasted mean summer discharge, meaning comparisons are less easily drawn with the monthly time step used in this study. However, the inability of both full year and stratified models to improve on the ZOF for Nurek inflows aligns closely with the findings of Schär et al. (2004) and Barlow and Tippett (2008) for the Vakhsh and nearby Murgab Rivers respectively. Comparable R^2 values are found using stratified models compared with the findings of Schär et al. (2004) ($R^2 = 0.85$) at a lead of one month for Toktogul ($R^2_{adj} = 0.80$) and Andijan ($R^2_{adj} = 0.77$) reservoirs. However, stratified model skill cannot compete with the mean summer flow forecast skill of Apel et al (2017) at any lead time or Schär et al. (2004) at a lead time longer than one month. Archer and Fowler (2008) found better model performance for April-June than April-September mean discharge for the River Jhelum, Pakistan. A similar shortening the target period may be beneficial in CA, focusing the forecast on the months during which snowmelt dominates the flow regime.

6.2 Climate mode based seasonal forecasting

6.2.1 Correlation analysis

Correlation analysis between NAO, Niño 3.4 and DMI with CA precipitation and temperature was undertaken. Previous studies have shown weak concurrent negative correlations between NAO and precipitation over parts of CA (Aizen et al., 2001, Syed et al., 2006). Results presented here generally agree with earlier research, showing insignificant negative correlations between NAO and precipitation over the same areas. Positive correlations were found between Niño 3.4 and CA precipitation confirming previous findings. The association is strongest during the winter half-year, consistent with Dai and Wigley (2001) who report positive annual precipitation anomalies of up to 50 mm over CA during El Niño phases. Syed et al. (2006) and Mariotti (2007) showed the strongest correlations occur during autumn and spring. October, November,

128

December and March have areas of significant positive correlations across parts of central and western Tajikistan (Figure 5.9). However, the statistically significant positive correlations found by Mariotti (2007) for south-east Kyrgyzstan did not emerge from this analysis. Only an insignificant positive correlation was detected over north-east Kyrgyzstan during winter months.

Significant negative correlations found here between NAO and January and February temperature across Tajikistan are stronger than those described by Syed et al. (2006) and Hurrell (1997). This could reflect methodological differences, as Syed et al. (2006) used a composite of temperature, rather than correlation analysis. Furthermore, the four-month period used by both studies could be masking the significant values found here during January and February only. El Niño typically results in positive winter temperature anomalies in CA (Hurrell, 1997; Syed et al. 2010), a finding corroborated by the significant positive correlations found during January-February across Kyrgyzstan (Figure 5.10). No previous studies have reported the strong, significant negative correlation during May across northern Kyrgyzstan and Kazakhstan. In mountainous northern Kyrgyzstan, May is often the wettest month. Hence, temperature anomalies during this period are important in determining whether precipitation falls in liquid or frozen form at higher elevations.

The non-parametric Spearman's Rank correlation was used to test associations between mean summer (Apr-Sep) inflows and preceding winter NAO, Niño 3.4 or DMI over the periods 1941-1980 and 2001-2010. Significant ($p < 0.05$) negative correlations between January-February NAO and following summer inflows to all three reservoirs during 1941-1980 is consistent with the modest negative correlations between NAO and precipitation during January-March across some parts of CA (Table 5.6). The result suggests that there could be a negative association between NAO and winter rainfall/ snow accumulation, leading to an indirect link between winter NAO and summer inflows.

Highly significant ($p < 0.01$) positive correlations between Nurek summer inflows and previous winter Niño 3.4 during 1941-1980 align closely with the positive correlations between winter precipitation and Niño 3.4 across Tajikistan (Table 5.6 and Figure 5.9). This suggests that Niño 3.4 has power to discriminate winter precipitation accumulations and thence summer inflows. Weaker, but still significant ($p < 0.01$)

correlations between winter Niño 3.4 and following summer inflows to Andijan and Toktogul again fits with the positive, though insignificant, correlations with precipitation over north eastern Kyrgyzstan during several winter months (Figure 5.9). Positive, but insignificant, correlations persist during the years 2001-2010 between winter Niño 3.4 and following summer inflows to all three reservoirs.

Overall, significant correlations between selected winter months Niño 3.4 and NAO to following summer reservoir inflows suggest potential predictability. In particular, forecasting Nurek summer inflows from winter Niño 3.4 appears most feasible (Table 5.6). Moreover, by using Niño 3.4 in November-December, it would be possible to issue an early outlook of summer inflows – well before the end of the winter snowpack accumulation period (October-March). Even so, these results must still be interpreted with care because of the large number of statistical tests performed and likelihood of Type I errors (i.e. false correlations). Given an error rate $p = 0.05$, approximately 4 significant results could have occurred by chance alone when correlating Niño 3.4 to following summer inflows. However, with 44 significant associations between Niño 3.4 and summer inflows actually detected the link is robust.

6.2.2 Composite forecasting

Composite analysis of data in the years 1941-1980 supports the findings of the correlation tests, with Niño 3.4 emerging as the most promising variable for predicting inflows, especially for Nurek (Figure 5.11). Differences of up to ±11% occur in summer inflows to Andijan following contrasting NAO phases, but differences were not statistically significant for any reservoir (Table 5.7 and Table 5.8). Inflows after positive NAO show greater differences with neutral conditions compared with negative NAO. The below average summer flows after positive NAO phases found here is consistent with Aizen et al. (2001) and the prior correlation analysis. Negative correlation between NAO and winter precipitation, as reported by Aizen et al. (2001) and Syed et al. (2006), would translate into less snowpack thence lower summer reservoir inflows as shown in Table 5.6.

Typically, El Niño conditions during winter are followed by higher summer inflows to all three reservoirs. The summer inflow anomaly for Nurek reservoir after a winter El Niño

can be as much as +37%. Figure 5.11 as well as Table 5.7 and Table 5.8 suggest a southwest to northeast gradient in the influence of Niño 3.4 index phase on discharge. Previous research suggests that the link between Pacific SSTs and CA rainfall anomalies is not direct, but via intervening atmospheric circulations over the Arabian Sea and Middle East (Mariotti, 2007; Barlow and Tippett, 2008). The gradient in predictability seen here could reflect depth of moisture penetration from the southwest. The Pamirs hinder moisture from reaching the Tien Shan, thereby reducing the connection between El Niño and precipitation in northwest CA (see section 6.2.5).

The influence of concurrent air temperature on summer inflows varies by reservoir. It might be expected that higher temperatures would be associated with greater meltwater and therefore, increased summer inflows, as seen for Nurek (Figure 5.12). However, the opposite occurs for Toktogul and Andijan reservoirs following neutral of El Niño conditions, with above average temperatures associated with below average inflows. Similar results were found by Archer and Fowler (2008) for the Indus basin, attributed to higher evaporative losses (Singh and Bengtsson, 2005). This is a plausible cause for both Toktogul and Andijan, as Nurek has a greater proportion of glacier-derived flow into Nurek compared with Toktogul and Andijan (Savoskul and Smakhtin, 2013). If all winter snow accumulation melts each year regardless of the temperature anomaly, summer runoff will be moisture rather than energy (temperature) limited. Research suggests that winter precipitation totals are a useful predictor of summer flow, particularly for the Syr Darya, giving weight to this hypothesis (Schär et al., 2004; Dixon and Wilby, 2016).

6.2.3 Stationarity of seasonal forecasting relationship at Nurek

Reconstruction of the discharge record at Komsomolabad (using a linear relationship with upstream flows at Garm, see section 4.3.2) allowed assessment of the stationarity of the Niño 3.4-Nurek inflow relationship during the years 1941-2016 (1993-1999 missing). The period 1981-2016 again showed increased (decreased) inflows to Nurek following El Niño (La Niña) conditions during October-March (Figure 5.13), but differences were statistically insignificant. This differs from the particularly strong association between ENSO and CA precipitation during 1980-2006 reported by Mariotti (2007). It is possible that missing data in this study (1993-1999) partly accounts for the

131

discrepancy. In particular, the very strong El Niño of 1997-1998 could not be included in the analysis because of missing data yet, according to CA-Water Info[28], coincided with greatest inflow to Nurek during the period 1981-2016. Moreover, it is conceivable that higher mean air temperatures in the latter period (Figure 2.6) have favoured proportionately greater contributions of glacier melt to total summer inflows, thereby weakening the apparent link to between Niño 3.4 and winter snowpack.

The statistically insignificant discriminatory power of Niño 3.4 during the later period could be related to the weakening of the ENSO-Indian Monsoon relationship in recent decades (Kumar et al., 1999 and Krishnaswamy et al., 2015). The weakening of the ENSO-Indian Monsoon relationship has been attributed to a warmer climate and a strengthening of IOD-Indian Monsoon relationship (Kumar et al., 1999 and Ashok et al., 2001). It is possible that the link between ENSO and CA is, like ENSO and the Indian Monsoon, becoming weaker and/or that it is being modulated by other climate modes such as the IOD. However, the strong, significant ($p < 0.01$), differences over the full record (1941-2016, 1993-1999 missing) and the physical mechanism provided give confidence in ENSO having discriminatory power in Nurek inflows.

Using November-December Niño 3.4 to forecast following summer inflows allows an early outlook forecast to be produced in January, before a second forecast could be issued in April based on October-March Niño 3.4. November- December Niño 3.4 shows greater discriminative power during the period 1941-1980, but less so during 1981-2016 and 1941-2016 as a whole. This pattern is replicated by the relative skill of each model for different validation periods, with November-December Niño 3.4 showing greater skill than October-March using the validation period 1941-1974 (Table 5.13). However, using the validation period 1974-2016 (1993-1999 missing) both forecasts yield equal skill according to the HHP (score of 0.51). It might be expected that the later information available from October-March Niño 3.4 would provide improved skill relative to November-December Niño 3.4. However, the HHP skill scores do not match this expectation. This could be explained by the strong relationship during November-December being diluted across October-March, therefore reducing forecast skill.

---

[28] http://www.cawater-info.net/amudarya/index_e.htm
Figure entitled "Dynamics of inflow to the reservoir during the growing and non-growing seasons"

6.2.4 Sensitivity testing

Significant differences ($p < 0.01$) were present between inflows to Nurek during warm versus cool summers when using Sari-Tash observed temperature during the period 1941-1980 (Table 5.7). However, when using Lyakhsh temperature to condition summer inflows no significant differences were present (Figure 5.14). It is possible that Sari-Tash, being at an elevation of 3155 m compared to Lyakhsh at 2100 m, is more representative of temperatures at higher altitudes where most snow and ice melt originates. Furthermore, even when using Sari-Tash observed summer temperature no significant differences occur in inflows for the period 1981-2016 (1993-1999) missing. Such sensitivity reduces confidence in the ability of temperature to discriminate summer inflows to Nurek reservoir.

The sensitivity of Nurek summer inflows (conditioned on prior November-December Niño 3.4) to the period used to define tercile boundaries was also assessed. Inflows following La Niña/El Niño show little sensitivity to tercile definition period (Figure 5.15). Inflows following neutral conditions are not as stable, with greater discriminatory power shown when using the period 1982-2010 to define tercile boundaries. Using the period 1982-2010, a season must exhibit a greater Niño 3.4 anomaly to be classed as La Niña/El Niño (-0.60/+0.62) compared with the period 1941-1980 (-0.46/+0.14). The change in discriminatory power, therefore, seems to be a result of the proportionately stronger La Niña/El Niño events in recent decades. This raises the boundary definition for La Niña/El Niño thereby only capturing relatively strong events from the earlier period, and hence improving discrimination.

6.2.5 Proposed physical mechanism

Confidence in statistical inference is strengthened when plausible physical explanations are offered (McGregor, 2017). Syed et al. (2006) suggested a direct influence of the warm ENSO phase on CA precipitation. However, later research proposes that an atmospheric moisture flux enters CA via the southwest, circulating clockwise around a large high pressure anomaly (Figure 6.2A) (Mariotti, 2007). In this model, moisture is advected across the Middle East mainly from the Arabian Sea. During December-January-February the moisture flux moves southwards, resulting in westerly transport of

moisture around 20°N and is only connected to CA via a secondary southerly flux (Mariotti, 2007). This mechanism has since been corroborated by several studies for both CA and adjacent regions (Barlow and tippet, 2008; Yadav et al., 2010; Cannon et al., 2017). Further evidence for a south-westerly moisture flux into CA during El Niño is provided by Figure 6.2B and Figure 6.2C which show positive correlations between Niño 3.4 and GPCP precipitation across the Middle East south west of CA. Negative correlations are clearly seen during September-November over the high pressure anomaly in the Indian Ocean (Figure 6.2B), with positive correlations pivoting clockwise around this high pressure anomaly. Whilst not as clear during March-May, positive correlations again occur north of the Gulf of Aden and strengthen over south-west CA (Figure 6.2C). This mechanism neatly explains the spatial and temporal variation in correlations presented here between Niño 3.4, precipitation and discharge.
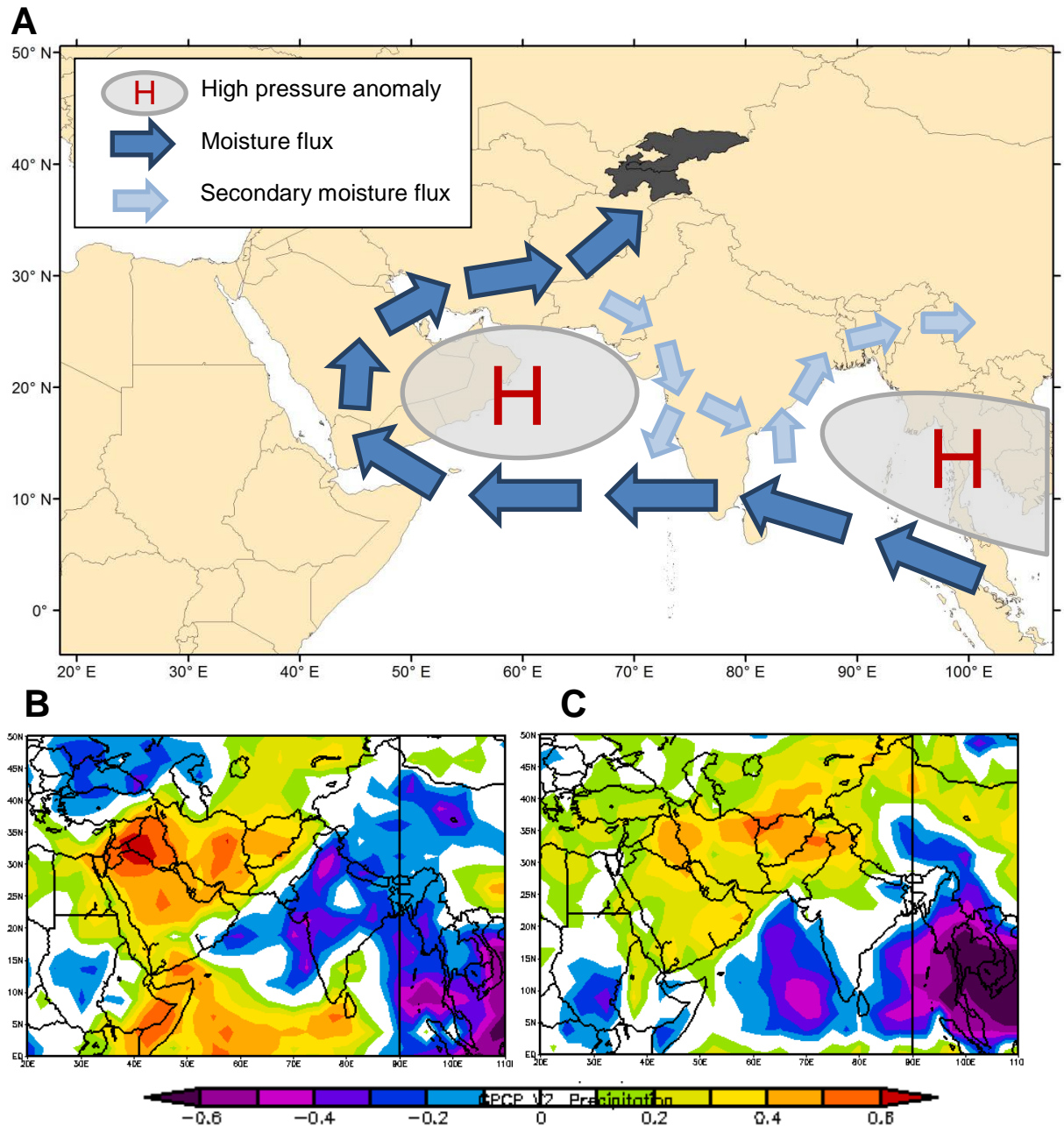
Figure 6.2: Model of enhanced moisture flux to Central Asia under El Niño conditions during September to November and March to May (Plate A, adapted from Mariotti, 2007). Linear correlation coefficient between Niño 3.4 and GPCP precipitation for the period 1979-2015 during September-November (Plate B) and March-May (Plate C).

6.3   Transferability of methodologies

6.3.1   Statistical forecasting model

NCEP precipitation shows significant positive correlations with M'Jara flow during the period 1952-1989 (Figure 5.16). Correlations are still significant for lead times up to three months. Strong, significant negative correlations between NCEP temperature and flow are likely due to them being in anti-phase. At lead times 1-3 months, averaging either precipitation or temperature reduces correlation strength (Figure 5.17). These contrast with findings for CA, likely due to the precipitation driven flow regime at M'Jara compared with the snow/ice melt driven regimes of CA.

Only at a lead time of one month were predictor variables added to the model beyond the dummy variable for month using the stepwise procedure. Again, this highlights the precipitation driven regime, lacking the snow accumulation which provided predictability at lead times of several months for CA reservoirs. Furthermore, Figure 5.18 shows modelled flows lag observed flows by one month resulting in a positive residual following a large negative residual. This is likely explained by the inclusion of antecedent flow in the regression model. This suggests that in flashy catchments, such as M'Jara, a time-step of less than (lag) one month is required to better capture observed flows. Whilst improved skill metrics are achieved by the model compared to ZOF at one-month lead time, the model cannot be described as functional with NSE 0.30. These results suggest that the statistical forecasting model approach is more reliant upon on slowly varying catchment properties such as the amount of snowpack or soil moisture status. Furthermore, the model may perform better in large basins (>10,000 km^2), with more damped flow regimes compared with smaller (M'Jara basin is 6,000km^2), precipitation dominated, flashier catchments.

6.3.2   Climate mode-based approach

The significant ($p < 0.01$) negative concurrent correlations found here between NAO and winter precipitation across northwest Morocco align closely with previous literature (Marshall et al., 2001). However, Marchane et al. (2016) report significant correlations at lead time one month for December, January and February in contrast to Figure 5.19. This discrepancy could be explained by the different time periods used (1952-1989 here compared with 1993-2012 by Marchane et al. 2016). Significant correlations between

Niño 3.4 and CA winter precipitation persisted to lead time three months. NAO has a more direct effect on Moroccan precipitation compared with the remote links between climate modes and CA precipitation (Hurrell et al., 2003). This direct link between NAO and precipitation in Morocco will likely reduce the lead time at which a skilful forecast can be made compared with CA. This is confirmed by Table 5.17, with significant correlations between NAO and M'Jara flow occurring only during the two months prior to the target flow period.

Of the two target flow periods, February-April mean shows greatest potential for conditioning. Significant differences were found between February-April flows following negative versus neutral and negative versus positive NAO during December-January Table 5.18. Figure 5.19 shows that large flow anomalies at M'Jara (>200 m^3s^{-1}) have a 38% likelihood following negative NAO, compared with 0% following neutral or positive NAO phases. Perhaps of greater utility to reservoir managers is that following positive NAO there is only a 15% likelihood of a positive flow anomaly and 38% chance of a negative anomaly in the band <100m^3s^{-1}. With this information action could be taken to conserve water in the reservoir reducing the impact if a large negative inflow anomaly occurs.

Compared with Niño 3.4 forecasts of Nurek inflows, significant differences in M'Jara flows can only be achieved by a forecast issued with zero lead time. Furthermore, according to KS tests, differences between exceedance probabilities are smaller for February-April M'Jara flows compared with summer inflows to Nurek (D-statistic of 0.62 compared with 0.85). Table 5.14 also shows the generally normal distribution of Nurek inflow anomalies following different Niño 3.4 phases, compared with the bi-modal distribution of M'Jara flows following negative NAO. This spread of possible flows at M'Jara of <100 to >200 m^3s^{-1} further reduces the utility of a forecast, as no flow anomaly category can be ruled out. Despite these limitations, a forecast of February-April mean flow at M'Jara could be of value for decision makers, particularly following positive NAO phases.

Overall, both approaches show greater potential for implementation in CA than in Morocco. The relationship between winter precipitation accumulation and summer reservoir inflows in CA (Schär et al., 2004 and Table 5.2) enables predictability at longer

lead times than for the study catchment in Morocco. Furthermore, the smaller, flashier catchment and hydrological regime of M'Jara lends itself to predictability at lead times of days rather than months. For flow forecasts at such lead times products such as the GLObal Flood Awareness System[29] could be used.

6.4   Operationalising a seasonal forecasting approach

Despite much effort, there still exists a gap between state-of-the-art research and the operational practices of seasonal hydrological forecasting (Pegano et al., 2014). This section will first review some of the barriers to operational implementation of seasonal forecasts in CA before assessing the potential for implementing the climate mode-based approach presented in the thesis. This approach was selected for operationalisation after consultation with reservoir operators in which it was clear this approach presented several benefits. Forecasts of mean summer inflows are the current standard in region, therefore continuing with this approach may improve the likelihood of forecast uptake. Furthermore, the longer lead time possible using the climate mode approach was desired to complement Tajik Hydromet's current regression-based forecast issued each April. Importantly, the approach is also relatively quick and straightforward to produce operationally.

One such issue between research and operational seasonal forecasting is perceived forecast value. Skill is a necessary ingredient in perceived forecast value, yet more forecast skill by itself does not imply more forecast value (Feldman and Ingram, 2009). Murphy (1993) stated that forecasts possess no intrinsic value, but instead acquire value through their capability to influence the decision making of forecast users (Thornes and Stephenson, 2001). Therefore, an increase in skill may not necessarily result in increased value if the forecast does not bring about improved decision making (e.g. due to lack of capacity in forecast interpretation, contractual obligation, economic concerns etc.) (Rayner et al., 2005; Crochemore et al., 2016). Still, the majority of seasonal flow forecasting literature focusses effort on improving the skill of outlooks, with relatively little attention given to the issues surrounding communication of forecasts (Demeritt et al., 2010).

---

[29] http://globalfloods.jrc.ec.europa.eu/

Investments in seasonal forecasting have often resulted in improved forecast services (e.g. more frequent updates). However, it is much less clear how, or if, these translate into improved water management decisions (Chiew et al., 2003; Werner et al., 2013). Hartmann et al. (2002) argued that forecasts should be evaluated for their "skill" at communicating their information content in ways that can be correctly interpreted both easily and reliably. Werner et al. (2013) found that only rarely have improved forecasts translated into greater forecast utilization by water management agencies. A possible reason for the focus on improving skill and reliability of seasonal forecasts as opposed to ease of communication and improved decision making is the difficulty in assessing these factors. Whilst skill can be measured via multiple, readily calculated metrics (e.g. Nash-Sutcliffe efficiency, Brier´s probability score), quantifying the communicability of a forecast is a more complex prospect.

Several institutional factors have been found to affect the use of novel forecasting techniques by decision makers. Rayner et al. (2005) assert that conservatism and complexity are the principal factors constraining the uptake of new forecasting approaches. The risk-averse system in which many institutions work results in managers relying on traditional methods to avoid the risk of public criticism (or worse) for using a new and unproven system if a wrong decision is made (Lemos, 2008). A need to adhere to industry standard procedures, such as government regulations regarding dam release for flood-control functions, can further reduce the role forecasts play in decision making (Georgakakos et al., 2005; Rayner et al., 2005). It is also important for decision makers within the organisation to understand and trust the forecast. This is a particular issue for probabilistic forecasts because some hydrologists and managers are not used to working with probabilities (Ludwig, 2009; Kirchhoff et al., 2013).

Mantua et al. (2008) identified a number of common characteristics of useful forecasts. Interactions between scientists and forecast users should be 'end-to-end' (i.e. utilising boundary organisations to translate and mediate between the producers and consumers of forecasts). A stable funding platform is critical to forecast uptake, as users must be confident that the program will continue to provide the required forecasts. In reality, most operational practice operates in a relatively *ad-hoc* nature with small adjustments

to forecast practices made when opportunities arise with the exception of larger/longer term projects (e.g. ECMWF, IRI, GLOFAS). Whilst procedures do exist, there is generally a lack of formal procedures across agencies for certifying/operationalising a new model/procedure (Mantua et al., 2008). This lack of structure results in the diverse forecasting approaches observed today, with no consistent agreement on the 'best' methods of how to validate forecast skill and, arguably more importantly, forecast value.

The climate mode forecasting approach was first developed before being presented to reservoir operators. After feedback on the required lead times and preferred presentation method, the approach was refined to the final form shown in Appendix B. Another approach would be to co-develop a seasonal forecasting system with suppliers and users of the outlook. Such an approach would have beyond the scope of the thesis. Furthermore, in data sparse situations like Tajikistan a pragmatic approach is required that is based on simple lookup tables and step by step procedures. Despite the relative simplicity of the approach developed, it is founded on sound exploratory analysis and physical reasoning. Due to these factors, the practical approach taken was deemed most likely for uptake by reservoir operators, hopefully resulting in improved decision making.

The forecast can be presented using a likelihood table, as in Table 5.14. This technique neatly presents the most likely summer inflow anomalies following La Niña/neutral/El Niño conditions during the previous winter. In addition, the table can be used to forecast the likelihood of more extreme conditions, conveying the spread of inflow anomalies following a particular phase. This forecasting approach gives some information regarding the uncertainty of the outlook, but inflows outside of these ranges could still occur. This could be caused by non-stationarity in the Niño 3.4-Nurek inflow relationship or because the relationship is not fully explained by the period of record used to define the association. After consultation with reservoir operatives, a forecast issued in January was deemed to complement existing forecast issued by Tajik Hydromet in April (using linear regression to relate winter observed precipitation to following summer inflows). Advanced technical notes provide more detailed information regarding the methodology and scientific judgements made throughout the model building process, in order to make the procedure as transparent as possible.

Despite the relative simplicity and pragmatism of this approach there are still a number of limitations. First, a judgement must be made by decision makers about what constitutes *practical* as opposed to *statistical* skill. For instance an HHP of 51-59% exceeds chance (33%) but whether this level of skill is useful partly depends on the appetite for risk of failure by the user. Second, flexible operating rules are a prerequisite of improved reservoir management linked to seasonal forecasts. Following consultation with reservoir operators in Tajikistan, it became apparent that this might not be the case. Operators seemingly had little flexibility in adjusting 'business as usual' operating procedure to utilise a seasonal forecast. Third, the forecast depends upon timely publication of the Niño 3.4 index. This reliance on a third party (in this instance the National Oceanic and Atmospheric Administration (NOAA)) could cause problems, as delays in index publication as well as decisions regarding funding, methodological changes or even termination of index publication is beyond the decision maker's control.

Forecast value is also partly determined by the active capacity of the reservoir – in the case of Nurek, this is just 4 km^3 of 10.5 km^3 total storage capacity. The active storage is fully utilised during all winters then replenished during most summers, with the reservoir regularly spilling most summers. In this case, the value of a forecast is drastically reduced if the reservoir is currently used at capacity every year, leaving no room for extra winter energy production during an El Niño winter. However, with the Central Asia South Asia Electricity Transmission and Trade Project[30] (CASA-1000) expected to be completed in 2018, a potential 1300 megawatts could be exported to Pakistan during summer months. This would increase demand on summer reservoir levels, meaning forecasts of summer inflows will still be beneficial to reservoir operation. Additionally, power production from the Rogun hydropower project is expected to begin in 2018. Seasonal forecasts could help plan reservoir filling and operations once commissioned. For example, if the winter prior to filling is La Niña, a more conservative approach could be taken to ensure downstream summer irrigation requirements can still be met.

The climate mode-based forecasting approach is not limited to only forecasting reservoir inflows. It is plausible that the approach could have skill at forecasting natural hazards (e.g. avalanches and mudflows) and be used to facilitate improved

---

[30] http://projects.worldbank.org/P145054?lang=en

transboundary water management (e.g. agree release strategies for certain severity forecasts). Possibly the greatest opportunity for development is forecasting of winter energy demand. Figure 5.10 shows significant ($p < 0.05$) correlations between Niño 3.4 and January/February temperatures across Tajikistan and Kyrgyzstan. Such a forecast could allow stockpiling of energy reserves (e.g. coal), planned imports of electricity (through CASA 1000) and even prepare emergency humanitarian relief in the event of a cold winter.

6.5   Summary

This chapter has contextualised the results of the thesis, discussing the findings in relation to the wider literature and within the thesis as a whole. First, results from the statistical forecasting approach were evaluated. Both research and operationally viable models were able to outperform the ZOF for Toktogul, Andijan and Kayrakkum reservoir inflows at lead times of one to four months (according to R^2_{adj}, NSE and RMSE skill metrics). Furthermore, stratifying models to forecast monthly inflows for only summer months (April-September) provides greater skill improvements over ZOF for lead times one, two and four months for Toktogul and Andijan. Stratified models incorporated NCEP precipitation with no averaging period at lead time three months, in contrast to full year models. Therefore, the improved skill of full year models is likely due to winter precipitation being better represented, a factor known to be important in forecasting summer inflows (Schär et al., 2004; Dixon and Wilby, 2016). All model types struggled to improve beyond the skill of the ZOF for Nurek inflows at all lead times investigated. It is possible that the relatively poor performance of precipitation predictors (Schiemann et al., 2008 and Table 3.4) as well as the increased influence of glacier melt reduce the forecast skill for Nurek reservoir (Savoskul and Smakhtin, 2013).

Of the climate modes investigated, Niño 3.4 showed the greatest potential for conditioning following summer inflows to Toktogul, Andijan and Nurek reservoirs. According to both KS and KW tests, significant ($p < 0.01$) differences occur in summer inflows following opposing phases of winter Niño 3.4 during the period 1941-1980. However, significant differences in Niño 3.4 conditioned Nurek inflows did not hold when using the period 1981-2016 (1993-1999 missing). It is conceivable that higher mean air temperatures in the latter period (Figure 2.6) have favoured proportionately greater

contributions from glacier (rather than winter snowpack) melt to total summer inflows, thereby weakening the apparent link to winter snowpack. Additionally, concurrent temperature showed potential for further conditioning inflows to Nurek reservoir ($p < 0.01$ for the period 1941-80). However, the temperature-inflow relationship did not hold when using Lyakhsh meteorological station rather than Sari-Tash. The higher elevation of Sari-Tash could be providing a better proxy for snowmelt. Nurek reservoir operatives would value an early outlook issued each January. Using November-December Niño 3.4 allows this and provides greater skill than October-March Niño 3.4 at forecasting following summer inflows to Nurek according to HHP. This increased skill derives from Niño 3.4 strength peaking during November-December, whereas using October-March Niño 3.4 dilutes the relationship. A physical mechanism was proposed for the statistical relationship, involving moisture being advected across the Middle East mainly from the Arabian Sea during El Niño conditions resulting in increased winter precipitation in CA.

The transferability of both forecasting approaches was then assessed via implementation for the M'Jara catchment, Morocco. NCEP precipitation was significantly ($p < 0.01$) positively correlated with monthly flow at M'Jara at lead times of up to two months for the period 1952-1989. However, predictors (in addition to monthly dummy) were only included in the model at lead time one month. Furthermore, although improving skill relative to the ZOF at one-month lead, the poor NSE of 0.30 is such that it cannot be deemed a functional model. The model struggles to depict the timing of peak flows, lagging observed flows by one month – likely the result of antecedent flow driving model skill. Overall, the difference in hydrological regime results in the approach performing poorly for the M'Jara catchment. Skill of CA inflows is driven by the relationship between winter snowpack accumulation and summer runoff. The flashier hydrology at M'Jara along with the minimal winter snow storage means skill is much reduced using this approach for the M'Jara catchment. Clearly, this assessment is based on only a single catchment, therefore more thorough investigation is required to test the wider validity of this assertion.

Significant differences ($p < 0.01$) were present between February-April inflows following opposing phases of December-January NAO for the period 1952-1989. For example, there is a 38% likelihood of February-April flows exceeding $200 m^3 s^{-1}$ following negative December-January NAO, compared with 0% following negative or neutral conditions.

However, seasonal forecasting skill is only provided by NAO with zero lead time (i.e. February-April forecast issued in February), limiting the forecasts value. Again, the difference in hydrological regime as well as direct impact of NAO compared to the more remote influence of ENSO on CA inflows reduces the potential in Morocco compared with CA.

After consultation with reservoir operators in Tajikistan it was clear that an early outlook forecast of mean summer reservoir inflows to Nurek issued in January would complement their current forecast which is provided by Tajik Hydromet each April. Using November-December Niño 3.4 allows a forecast of following summer inflows to be issued in January with a success rate of 51-59% depending on period of record (assessed via HHP). Further consultation is needed to establish whether this level of statistical skill is of practical significance for operational decision making, and/or the extent to which forecast value is limited by inflexible operating rules. The skill of the existing system operated by Tajik Hydromet should also be assessed using HHP, allowing comparison between approaches. Furthermore, Nurek's hydropower potential is fully used every winter and water is spilled nearly every summer, severely limiting the utility of a forecast to manage winter energy shortages. However, forecasts could be used to improve dam safety during large inflows and for scheduling maintenance, or Rogun filling. Furthermore, the approach presented here could be expanded to forecast natural hazards (e.g. avalanches and mudflows) and help facilitate improved transboundary water management (e.g. by negotiating different operating rules depending on the seasonal outlook).

The final Chapter begins with a review of the thesis, identifying the key findings of each section of the thesis. The major findings of the thesis are then summarised, beginning with the collation and quality assessment of public domain hydro-climatic information. The two forecasting approaches are then reviewed, discussing circumstances in which the models perform well/poorly as well as suggesting avenues for further research which could lead to improvements in model performance. Transferability of approaches is then reviewed using the M'Jara catchment in Morocco. The Chapter concludes with suggestions for extending the climate mode approach, for example to forecasting winter energy demand.

7 Conclusions and further research

This thesis developed information streams and model structures to facilitate reservoir inflow forecasting in data sparse regions using public domain predictors. Chapter 1 set out the motivations for the study. Winter electricity rationing regularly occurs in both Kyrgyzstan and Tajikistan with serious consequences for human well-being. As most electricity is produced by Toktogul and Nurek reservoirs respectively, the energy crises are directly linked to management of these critical facilities. Previous research was then reviewed in Chapter 2. This highlighted the fragmentary hydro-climatological gauge network in the region, making near real-time satellite and re-analysis products appealing. Chapter 3 introduced the reservoir sites, before collating and selecting public domain hydro-climatic data to address the aim and objectives of the thesis.

The methodologies used in the thesis were described in Chapter 4. A regression approach was adopted as it is transparent, tractable and transferable. The climate mode-based forecasting approach used 0.33 and 0.66 index percentiles to condition inflows, selected to allow a roughly equal number of seasons in each category. Chapter 5 showed that, overall, both research grade and operational models were able to outperform the ZOF for Toktogul, Andijan and Kayrakkum for all lead times using most skill metrics. Stratifying models to forecast only summer inflows further improved skill at lead times 1, 2 and 4 months for Toktogul and Andijan reservoirs. Significant differences ($p < 0.01$) were present between summer inflows to all three reservoirs conditioned on prior October-March Niño 3.4, as well as for Nurek inflows conditioned on concurrent temperature for the years 1941-1980. Both approaches showed greater skill at longer lead times for CA than for the test site in North Africa.

These results were then discussed with respect to prior literature in Chapter 6. Of the climate modes assessed, Niño 3.4 showed greatest power at discriminating summer reservoir inflows in CA. This is consistent with latest understanding of the atmospheric processes involved, by which a moisture flux from the Arabian Sea moves clockwise around a block of high pressure centred over the Indian Ocean (Mariotti, 2007). This final Chapter summarises the major findings of the thesis and offers suggestions for further research.

7.1 Headline findings

- Both TRMM and NCEP precipitation products are significantly correlated with monthly observed values at Naryn and Uzgen (Syr Darya basin/Tien Shan) but not at Sari-Tash (Amu Darya basin/Pamirs).

- Near real-time TRMM overestimates precipitation several fold during both autumn and spring, attributed to issues in defining precipitation phase.

- Statistical models perform better at forecasting inflows in the Tien Shan (Naryn and Andijan) than the Pamir (Nurek).

- Stratifying models to forecast only summer inflows improved model skill relative to the Zero Order Forecast.

- Significant correlations ($p < 0.01$) were present between most winter monthly and bi-monthly Niño 3.4 with following summer mean inflows to Toktogul, Andijan and Nurek during the years 1941-1980.

- According to both Kolmogorov-Smirnov and Kruskal-Wallis tests, significant ($p < 0.01$) differences occur in summer inflows to all three reservoirs following opposing phases of winter Niño 3.4 during the period 1941-1980.

- Significant differences for Nurek inflows did not hold, however, when assessed the period 1981-2016 (1993-1999 missing).

- Using November-December Niño 3.4 allowed an early outlook forecast whilst providing greater skill than October-March Niño 3.4 at forecasting following summer inflows to Nurek according to the Heidke Hit Proportion (51-59% depending on period of record).

- A physical mechanism for the statistical relationships was proposed (based on Mariotti, 2007) in which moisture is advected across the Middle East mainly from the Arabian Sea during El Niño conditions.

7.2 Further avenues for research

- The cause(s) of recent weakening of the Niño 3.4-Nurek inflow relationship could be explored via perturbing Ocean temperatures within climate model experiments and examining the response of winter precipitation across CA.

- Procurement of recent inflow data to Toktogul and Andijan reservoirs would allow the stationarity of Niño 3.4-CA inflow relationship to be evaluated more

146

thoroughly, and should be a priority to assess whether the weakening relationship at Nurek is representative of the wider region.

- A key research question within the region is whether 'peak water' (or the tipping point) has yet been reached. An observation-based study of water balance components including the relative contribution of glacier melt to river flows is, therefore, a key research priority (Unger-Shayesteh et al., 2013). This would no doubt require significant endeavour, with sustained funding as well as cooperation amongst government agencies and between nation states.

- The ability of other regional climate drivers (such as the Indian monsoon) to further condition summer inflows should be explored, potentially reducing the inflow range of seasonal forecasts.

- Expansion of the climate mode approach to forecast winter energy demand should be explored. This could be used to stock pile other energy supplies (e.g. coal) and/or pre-prepare aid agencies to help reduce the impact of the all to frequent winter energy crises.

7.3 Final remarks

The approaches presented in this thesis are low cost, require relatively limited technical or computing capacity, and can be applied to remotely sensed products – all considerations in data sparse, low capacity regions. The climate mode forecasting approach shows great potential for operationalisation in CA. Despite the limited data and technical capacity there is need for scientific rigour in the hunt for robust, stationary predictability. This must be matched with pragmatism about how such relationships and forecasting tools might be operationalised. The research has been translated into easily interpreted look up tables alongside step-by-step guidance notes, enabling uptake by reservoir operators.

8 References

Abdelbasset, M., Abderrahim, L., Abdel_Ali, C., Abdellah, B., Lahcen, B. and Laila, B. (2015). Inegration of GID and HEC-RAS in Floods Modeling of the Ouergha River, Northern Morocco. *European Scientific Journal*, **11**, 196-204.

Abdullaev, I. and Rakhmatullaev, S. (2013). Transformation of water management in Central Asia: from state-centric, hydraulic mission to socio-political control. *Environmental Earth Sciences*, **73**, 849–861.

Abrahart, R.J., Anctil, F., Coulibaly, P., Dawson, C.W., Mount, N.J., See, L.M., Shamseldin, A.Y., Solomatine, D.P., Toth, E. and Wilby, R.L., 2012. Two decades of anarchy? Emerging themes and outstanding challenges for neural network river forecasting. *Progress in Physical Geography*, **36**, 4, 480–513.

Adler, R.F. Huffman, G.J., Chang, A., Ferraro, R., Xie, P.P., Janowiak, J., Rudolf, B., Schneider, U., Curtis, S., Bolvin, D., Gruber, A., Susskind, J., Arkin, P. and Nelkin, E. (2003). The Version-2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979–Present). *Journal of Hydrometeorology*, **4**, 6, 1147–1167.

Agaltseva, N.A., Borovikova, L.N. and Konovalov, V.G. (1997). Automated system of runoff forecasting for the Amudarya River basin. *Destructive Water: Water-Caused Natural Proceedings of Disasters, their Abatement and Control (Conference proceedings)*. IAHS Publ. no. 239.

Auston, P., Mackay, A., Palagushkina, O. and Leng, M. (2007). A high-resolution diatom-inferred palaeoconductivity and lake level record of the Aral Sea for the last 1600 yr. *Quaternary Research*, **67**, 383-393.

Aizen, E.M., Aizen, V.B., Melack, J.M., Nakamura, T. & Ohta, T. (2001). Precipitation and atmospheric circulation patterns at mid-latitudes of Asia. *International Journal of Climatology*, **21**, 5, 535-556.

Aizen, V., Aizen, E., Melack, J. & Martma, T. (1996). Isotopic measurements of precipitation on central Asian glaciers (southeastern Tibet, northern Himalayas, central Tien Shan). *Journal of Geophysical Research: Atmospheres (1984–2012), **101**, 4, 9185-9196.

Aizen, V.B., Aizen, E.M and Melack, J.M. (1995). Climate, snow cover, glaciers and runoff in the Tien Shan, Central Asia. *Journal of the American Water Resources Association,* **31**, 6, 1113-1129.

Aizen, V.B., Mayewski, P.A., Aizen, E.M., Joswiak, D.R., Surazakov, A.B., Kaspari, S. Grigholm, B., Krachler, M., Handley, M. and Finaev, A. (2009). Stable-isotope and trace element time series from Fedchenko glacier (Pamirs) snow/firn cores. *Journal of glaciology*, **55**, 275-291.

Alfieri, L., Burek, P., Dutra, E., Krzeminski, B., Muraro, D., Thielen, J. and Pappenberger, F. (2013). GloFAS – global ensemble streamflow forecasting and flood early warning. *Hydrological and Earth System Sciences*, **17**, 1161-1175.

Antelava, N. (2008). *Winter shortages fuel Tajik anger*. Available: http://news.bbc.co.uk/1/hi/world/asia-pacific/7239279.stm. Last accessed 16/10/2017.

Apel, H., Abdykerimova, Z., Agalhanova, M., Baimaganbetov, A., Gavrilenko, N., Gerlitz, L., Kalashnikova, O., Unger-Shayesteh, K., Vorogushyn, S. and Gafurov, A. (2017). Statistical forecast of seasonal discharge in Central Asia for water resources management: development of a generic linear modelling tool for operational use. *Hydrology and earth Systems Sciences Discussions*, doi: 0.5194/hess-2017-340.

Archer, D.R. and Fowler, H.J. (2008). Using meteorological data to forecast seasonal runoff on the River Jhelum, Pakistan. *Journal of Hydrology*, **361**, 10-23.

Armstrong, R.L. (2010). *The Glaciers of the Hindu Kush-Himalayan Region: A Summary of the Science Regarding Glacier Melt/Retreat in the Himalayan, Hindu Kush, Karakoram, Pamir and Tien Shan Mountain Ranges.* International Centre for Integrated Mountain Development (ICIMOD), Kathmandu.

Artan, G., Gadain, H., Smith, J.L., Asante, K., Bandaragoda, C.J. & Verdin, J.P. (2007). Adequacy of satellite derived rainfall data for stream flow modeling. *Natural Hazards,* **43**, 2, 167-185.

Arushev, M.L., Tsarev, B.K. and Shetinnikov, A.S. (1977). Kvoprosu ovliyanii Indiiskogo mussona na regim lednikov Pamira [The question about influence of Indian monsoon on the regime of Pamir's glaciers]. Mater. Glyatsiol. Issled., **36**, 145–148. [In Russian with English summary.]

Ashok, K., Guan, Z. and Yamagata, T. (2001). Impact of the Indian Ocean dipole on the relationship between the Indian Monsoon rainfall and ENSO. *Geophysical Research Letters*, **28**, 4499–4502.

Ashok, K., Guan, Z., Saji, N.H. and Yamagata, T. (2004). Individual and Combined Influences of ENSO and the Indian Ocean Dipole on the Indian Summer Monsoon. *Journal of Climate*, **17**, 3141-3155.

Asian Infrastructure Investment Bank. (2017). *Nurek Hydropower Rehabilitation Project.* Available:
https://www.aiib.org/en/projects/approved/2017/_download/tajikistan/document/document_nurek-hydropower-rehabilitation-project.pdf. Last accessed 16/10/2017.

Balthrop, C. and Hossain, F. (2010). Short note: A review of state of the art on treaties in relation to management of transboundary flooding in international river basins and the Global Precipitation Measurement mission. *Water Policy*, **12**, 635-640.

Barlow, M. and Hoell, A. (2015). Drought in the Middle East and central-southwest Asia during winter 2013/14. *Supplement to the Bulletin of the American Meteorological Society*, **96**, 12, 71-76.

Barlow, M., Cullen, H. and Lyon, B. (2002). Drought in Central and Southwest Asia: La Niña, the Warm Pool, and Indian Ocean Precipitation. *Journal of Climate*, **15,** 7, 697-700.

Barlow, M.A. and Tippett M.K. (2008). Variability and Predictability of Central Asia River Flows: Antecedent Winter Precipitation and Large-Scale Teleconnections. *Journal of Hydrometeorology*, **9**, 1334-1349.

Barnston, A.G., Chelliah, M. and Goldenberg, S.B. (1997). Documentation of a highly ENSO-related SST region in the equatorial Pacific. *Atmosphere-Ocean*, **35**, 367-383.

Baumgartner, M.F., Spreafico, M. and Weiss, H.W. (2000). Operational snowmelt runoff forecasting in the Central Asian mountains. *Remote Sensing and Hydrology*, 66-71.

Betts, A.K., Zhao, M., Dirmeyer, P.A. and Beljaars, A.C.M. (2006). Comparison of ERA40 and NCEP/DOE near-surface data sets with other ISLSCP-II data sets. *Journal of Geophysical Research*, **111**, D22, doi:10.1029/2006JD007174.

Birkinshaw, S.J., O'Donnell, G.M., Moore, P., Kilsby, C.G., Fowler, H.J. and Berry, P.A.M. (2010). Using satellite altimetry data to augment flow estimation techniques on the Mekong River. *Hydrological Processes*, **24**, 3811–3825.

Bolch, T. (2007). Climate change and glacier retreat in northern Tien Shan (Kazakhstan/Kyrgyzstan) using remote sensing data. *Global and Planetary Change,* **56**, 1-12.

Boomer, I., Aladin, N., Plotnikov, I. and Whatley, R. (2000). The palaeolimnology of the Aral Sea: a review. *Quaternary Science Reviews*, **19**, 1259-1278.

Bosilovich, M.G., Chen, J., Robertson, F.R. and Adler, R.F. (2008). Evaluation of global precipitation in reanalyses. *Journal of Applied Meteorology and Climatology*, **47**, 2279–2299.

Bothe, O., Fraedrich, K. & Zhu, X. (2012). Precipitation climate of Central Asia and the large-scale atmospheric circulation. *Theoretical and Applied Climatology*, **108**, 3-4, 345-354.

Bowman, K.P. (2005). Comparison of TRMM precipitation retrievals with rain gauge data from ocean buoys. *Journal of Climate*, **18**, 178-190.

Breu, T., Hurni, H. & Stucki, A.W. (2003). *The Tajik Pamirs: Challenges of sustainable development in an isolated mountain region*. Geographica Bernensia.

Cannon, F., Carvalho, L.M.V., Jones, C., Hoell, A., Norris, J., Kiladis, G.N. and Tahir, A.A. (2017). The influence of tropical forcing on extreme winter precipitation in the western Himalaya. *Climate Dynamics*, **48**, 1213–1232.

CAWaterInfo. (2017a). *Reservoirs in Kyrgyzstan*. Available: http://www.cawater-info.net/bk/1-1-1-1-3-kg_e.htm. Last accessed 18/07/2017.

CAWaterInfo. (2017b). *Reservoirs in Tajikistan*. Available: http://www.cawater-info.net/bk/1-1-1-1-3-tj_e.htm. Last accessed 18/07/2017.

Cayan, D.R., Redmond, K.T. and Riddle, L.T. (1999). ENSO and Hydrologic Extremes in the Western United States. *Journal of Climate*, **12**, 2881–2893.

Chanysheva, S.G., Subbotina, O.I., Petrov, U.V., Egamberdiyev, K.T., Aizenshtat, B.A. & Leukhina, G.N. (1995). *Variability of the Central Asian Climate*, Central Asian Hydrometeorological Research Institute: Tashkent, (in Russian).

Chiew, F.H.S. and McMahon, T.A. (2002). Global ENSO-streamflow teleconnection, streamflow forecasting and interannual variability. *Hydrological Sciences Journal*, **47**, 3, 505-522.

Chiew, F.H.S., Piechota, T.C., Dracup, J.A. and McMahon, T.A. (1998). El Nino/Southern Oscillation and Australian rainfall, streamflow and drought: Links and potential for forecasting . *Journal of Hydrology*, **204**, 138-149.

Chiew, F.H.S., Zhou, S.L., and McMahon, T.A. (2003). Use of seasonal streamflow forecasts in water resources management. *Journal of Hydrology*, **270**, 135–144.

Clausen, T.J., Fuggle, R., Giovannetti, F., Helland-Hansen, E. and Todini, E. (2014). *Rogun Hydropower Project: Final Report of the Environmental and Social Panel of Experts.* Available: https://www.worldbank.org/en/country/tajikistan/brief/final-reports-related-to-the-proposed-rogun-hpp. Last Accessed 04/04/2017.

Climate Data Guide. (2017). *NCEP/NCAR (R1): An Overview*. Available: https://climatedataguide.ucar.edu/climate-data/ncep-ncar-r1-overview. Last accessed 08/07/2017.

Cloke, H. L. and Pappenberger, F. (2009) Ensemble flood forecasting: a review. *Journal of Hydrology*, **375**, 3-4, 613-626.

Collischonn, B. and Pante, A.R. (2011). TRMM-forced rainfall—runoff modelling for water management purposes in small ungauged basins. *IAHS-AISH publication*. ISSN-0144-7815.

Condom, T., Rau, P. & Espinoza, J.C. (2011). Correction of TRMM 3B43 monthly precipitation data over the mountainous areas of Peru during the period 1998–2007. *Hydrological Processes*, **25**, 12, 1924-1933.

Conrad, C., Schonbrodt-Stitt, S., Low, F., Sorokin, D. and Paeth, H. (2016). Cropping Intensity in the Aral Sea Basin and Its Dependency from the Runoff Formation 2000–2012. *Remote Sensing*, **8**, 8, DOI: 10.3390/rs8080630.

Crochemore, L., Ramos, M.H., Pappenberger, F., van Andel, S.J., and Wood, A.W. (2016). An Experiment on Risk-Based Decision-Making in Water Management Using Monthly Probabilistic Forecasts. *Bulletin of the American Meteorological Society*, **97**, 541–551.

Cruz-Del Rosario, T. (2009). Risky riparianism: cooperative water governance in Central Asia. *Australian Journal of International Affairs*, **63**, 3, 404-415.

Cullen, H.M. and deMonocal P.B. (2000). North Atlantic Influence on Trigris–Euphrates Streamflow. *International Journal of Climatology*. **20**, 853–863.

Cullen, H.M., Kaplan, A., Arkin, P.A. and Demenocal, P.B. (2002). Impact of the north Atlantic Oscillation on Middle Eastern Climate and Streamflow. *Climatic Change*. **55**, 315–338.

Dai, A. and wigley, T.M.L. (2000). Global Patterns of ENSO-induced Precipitation, *Geophysical Research Letters*, **27**, 9, 1283-1286.

Daly, C., Gibson, W.P., Taylor, G.H., Doggett, M.K. and Smith J.I. (2007). Observer Bias in Daily Precipitation Measurements at United States Cooperative Network Stations. *Bulletin of the American Meteorological society*. **88**, 6, 899-912.

Dawson, C.W, Abrahart, R.J. and See, L.M. (2007). HydroTest: A web-based toolbox of evaluation metrics for the standardised assessment of hydrological forecasts. *Environmental Modelling & Software*, **22**, 7, 1034-1052.

Demeritt, D., Cloke, H, Pappenberger F., Thielen, J., Bertholmes, J. and Helena Ramos, M. (2007). Ensemble predictions and perceptions of risk, uncertainty, and error in flood forecasting. *Environmental Hazards*, **7**, 115-127.

Demeritt, D., Nobert, S., Cloke, H. and Pappenberger, F. (2010). Challenges in communicating and using ensembles in operational flood forecasting. *Meteorological Applications,* **17**, 2, 209-222.

Dixon, S.G. & Wilby, R.L. (2016). Forecasting reservoir inflows using remotely sensed precipitation estimates: a pilot study for the River Naryn, Kyrgyzstan. *Hydrological Sciences Journal*, **61**, 1, 107-122.

Dixon, S.G. and Wilby, R.L. (2017). Potential for seasonal forecasting of reservoir inflows in Central Asia. *Environmental Research Letters*, under revision.

Ebert, E.E., Janowiak, J.E. and Kidd, C. (2007). Comparison of Near-Real-Time Precipitation Estimates from Satellite Observations and Numerical Models. *American Meteorological Society*, **88**, 47-64.

ECMWF. (2017). *Seasonal forecasts user guide*. Available: https://www.ecmwf.int/en/forecasts/documentation-and-support/long-range/seasonal-forecast-documentation/user-guide/introduction. Last accessed 27/06/2017.

Emerton, R., Cloke, H. L., Stephens, E. M., Zsoter, E., Woolnough, S. J. and Pappenberger, F. (2017). Complex picture for likelihood of ENSO-driven flood hazard. *Nature Communications*, **8**, doi:10.1038/ncomms14796.

Fan, F.M., Collischonn, W., Quiroz, K.J., Sorribas, M.V., Buarque, D.C. and Siqueira, V.A. (2016). Flood forecasting on the Tocantins River using ensemble rainfall forecasts and real-time satellite rainfall estimates. *Flood Risk Management*, **9**, 3, 278–288.

Fekete, B.M., Vorosmarty, C.J., Roads, J.O. and Willmott, C.J. (2003). Uncertainties in Precipitation and Their Impacts on Runoff Estimates. *Journal of Climate*, **17**, 294-304.

Feldman, D.L. and Ingram, H.M. (2009). Making Science Useful to Decision Makers: Climate Forecasts, Water Management, and Knowledge Networks. *Weather, Climate and Society*, **1**, 9-21.

Fleming, S.W. and Dahlke, H.E. (2014). Parabolic northern-hemisphere river flow teleconnections to El Niño-Southern Oscillation and the Arctic Oscillation. *Environmental Research Letters*. **9**, doi:10.1088/1748-9326/9/10/104007.

Fowler, T.L., Jenson, T.L. and Brown, B.G. (2012*). Introduction to forecast verification* [PDF document]. Available: http://www.dtcenter.org/met/users/docs/presentations/WRF_Users_2012.pdf.

Gafurov, A. (2010). Water Balance Modeling Using Remote Sensing Information - Focus on Central Asia. *PhD Thesis*. 16-19.

Gan, R., Luo, Y., Zuo, Q. and Sun, L. (2015). Effects of projected climate change on the glacier and runoff generation in the Naryn River Basin, Central Asia. *Journal of Hydrology*. **523**, 240-251.

Gebremichael, M., Anagnostou, E.N. and Bitew, M.M. (2010). Critical Steps for Continuing Advancement of Satellite Rainfall Applications for Surface Hydrology in the Nile River Basin. *Journal of the American Water Resources Association*, **46**, 361-366.

Georgakakos, K.P., Graham, N.E., Carpenter, T.M. and Yao, H. (2005). Integrating climate-hydrology forecasts and multi-objective reservoir management for northern California. *Eos, Transactions American Geophysical Union*, **86**, 122-127.

Gerlitz, L., Vorogushyn, S., Apel, H., Gafurov, A., Unger-Shayesteh, K. and Merz, B. (2016). A statistically based seasonal precipitation forecast model with automatic predictor selection and its application to central and south Asia. *Hydrology and Earth System Sciences*, **20**, 4605-4623.

Guo, H., Chen, S., Bao, A., Hu, B., Gebregiorgis, A.S., Xue, X. and Zhang, X. (2015). Inter-Comparison of High-Resolution Satellite Precipitation Products over Central Asia. *Remote Sensing*, **7**, 6, doi:10.3390/rs70607181.

Guo, L. and Li, L. (2015). Variation of the proportion of precipitation occurring as snow in the Tian Shan Mountains, China. *International Journal of Climatology*, **35**, 1379-1393.

Gupta, H.V., Wagener, T. & Liu, Y. (2008). Reconciling theory with observations: elements of a diagnostic approach to model evaluation. *Hydrological Processes*, **22**, 18, 3802-3813.

Hagg, W., Hoezle, M., Wagner, S., Mayr, E. and Klose, Z. (2013). Glacier and runoff changes in the Rukhk catchment, upper Amu-Darya basin until 2050. *Global and Planetary Change*, **110**, 62-73.

Hamlet, A.F. and Lettenmaier, D.P. (1999). Columbia River Streamflow Forecasting Based on ENSO and PDO Climate Signals. *Journal of Water Resources Planning and Management*, **125**, 6, 333-341.

Hamlet, A.F., Huppert, D.D. and Lettenmaier, D.P. (2002). Economic Value of Long-Lead Stream-Flow Forecasts for Columbia River Hydropower. *Journal of Water Resources Planning and Management*, **128**, 2, 91-101.

Harding, L. (2010). Kyrgyzstan violence: UN official accuses outside groups of planning attacks. Available: https://www.theguardian.com/world/2010/jun/16/kyrgyzstan-violence-un-accuses-outsiders. Last accessed 28/05/2017.

Havenith, H.B., Torgoev, I., Torgoev, A., Strom, A., Xu, Y., Fernandez-Steeger, T. (2015). The Kambarata 2 blast-fill dam, Kyrgyz Republic: blast event, geophysical monitoring and dam structure modelling. *Geoenvironmental Disaster,* **2**, doi: 10.1186/s40677-015-0021-x.

Hoell, A., Barlow, m., Cannon, F. and Xu, T. (2017). Oceanic Origins of Historical Southwest Asia Precipitation During the Boreal Cold Season. *Journal of Climate*, doi: 10.1175/JCLI-D-16-0519.1.

Hoell, A., Funk, C. and Barlow, M. (2014). La Niña diversity and Northwest Indian Ocean Rim teleconnections. *Climate Dynamics*, *43*, 2707-2724.

Hoerling, M. and Kumar, A. (2003). The perfect ocean for drought. *Science*, **299**, 691–694.

Hossain, F. and Huffman, G.J. (2008). Investigating Error Metrics for Satellite Rainfall Data at Hydrologically Relevant Scales. *Journal of Hydrometeorology*, **9**, 563-575.

Hossain, F., Katiyar, N., Hong, Y. & Wolf, A. (2007). The emerging role of satellite rainfall data in improving the hydro-political situation of flood monitoring in the under-developed regions of the world. *Natural Hazards*, **43**, 2, 199-210.

Hu, Z., Hu, Q., Zhang, C., Chen, X. and Qingxiang, L. (2016). Evaluation of reanalysis, spatially interpolated and satellite remotely sensed precipitation data sets in Central Asia. *Journal of Geophysical Research*, **121**, 10, 5648–5663.

Huang, W., Chen, F.H., Feng, S., Chen, J.H. and Zhang, X.J. (2013). Interannual precipitation variations in the mid-latitude Asia and their association with large scale atmospheric circulation. *Chinese Science Bulletin*, **58**, 3963–3968

Huffman, G.J. & Bolvin, D.T. (2013). *TRMM and other data precipitation data set documentation*. NASA, Greenbelt, USA, 1-40.

Huffman, G.J. & Bolvin, D.T. (2017). *TRMM and other data precipitation data set documentation*. NASA, Greenbelt, USA, 1-40.

Huffman, G.J., Adler, R.F., Bolvin, D.T. and Nelkin, E.J. (2010). The TRMM Multi-Satellite Precipitation Analysis (TMPA). In: Gebremichael, M. and Hossain, F. *Satellite Rainfall Applications for Surface Hydrology*. Springer Science & Business Media. 3-22.

Huffman, G.J., Bolvin, D.T., Nelkin, E.J., Wolff, D.B., Adler, R.F., Gu, G., Hong, Y., Bowman, K.P. & Stocker, E.F. (2007). The TRMM multisatellite precipitation analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *Journal of Hydrometeorology*, **8**, 1, 38-55.

Hurrell, J., Kushnir, Y., Ottersen, G., Visbeck, M (2003). *The North Atlantic Oscillation: Climatic Significance and Environmental Impact.* Washington: American Geophysical Union. 1-35.

Hurrell, J.W. (1995). Decadal Trends in the North Atlantic Oscillation: Regional Temperatures and Precipitation. *Science*, **269**, 676-679.

Hurrell, J.W. and NCAR Staff (Eds). (2016). *The Climate Data Guide: Hurrell North Atlantic Oscillation (NAO) Index (station-based).* Available: https://climatedataguide.ucar.edu/climate-data/hurrell-north-atlantic-oscillation-nao-index-station-based. Last accessed 14/12/2016.

Hydro World. (2010). *Kyrgyzstan launches Kambarata 2 hydropower plant.* Available: http://www.hydroworld.com/articles/2010/09/kyrgyzstan-launches.html. Last accessed 07/07/2017.

Hydro World. (2016*). Work begins on Tajikistan's 3,600-MW Rogun hydroelectric plant.* Available: http://www.hydroworld.com/articles/2016/10/work-begins-on-tajikistan-s-3-600-mw-rogun-hydroelectric-plant.html. Last accessed 07/07/2017.

Hydro World. (2017*). IDA announces financing for rehab of Tajikistan's 3,000-MW Nurek hydropower plant.* Available: http://www.hydroworld.com/articles/2017/05/ida-announces-financing-for-rehab-of-tajikistan-s-3-000-mw-nurek-hydropower-plant.html. Last accessed 18/07/2017.

Immerzeel, W.W., van Beek, L.P.H. and Bierkens, M.F.P. (2010). Climate Change Will Affect the Asian Water Towers. *Science*, **328**, 1382-1385.

Inagamova, S. T. Mukhtarov, S. & Muchtarov S. (2002). *Characteristics of Synoptic Processes of Central Asia (in Russian)*, Central Asian Hydrometeorology Research Institute, Tashkent, Uzbekistan.

International Crisis Group. (2014). *Water Pressures in Central Asia*. Available: https://www.crisisgroup.org/europe-central-asia/central-asia/water-pressures-central-asia. Last accessed 16/10/2017.

International Research Institute. (2013). *Descriptions of the IRI Climate Forecast Verification Scores.* Available: http://iri.columbia.edu/wp-content/uploads/2013/07/scoredescriptions.pdf. Last accessed 18/10/2017.

IPCC. (2013). Climate Change 2013: *The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J.

Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom.

Ismaiylov, G.K., Fedorov, V.M. and Sadati Nezhad, S.D. (2007). Assessment of possible anthropogenic changes in the runoff of the Syr Darya River on the basis of a mathematical model. *Water Resources*, **34**, 4, 359-371.

Ji, X. & Chen, Y. (2012). Characterizing spatial patterns of precipitation based on corrected TRMM 3B43 data over the mid Tianshan Mountains of China. *Journal of Mountain Science*, **9**, 5, 628-645.

Ji, X. and Yu, Y. (2013). The influence of precipitation and temperature input schemes on hydrological simulations of a snow and glacier melt dominated basin in Northwest China. *Hydrology and Earth System Sciences Discussion Paper*, 10, 807-853.

Jones, P.D., Lister, D.H. and Li, Q. (2008). Urbanization effects in large-scale temperature records, with an emphasis on China. *Journal of Geophysical Research*. **113**, doi:10.1029/2008JD009916.

Kalnay, E., Kanamitsu, M., Kirtler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K. C., Ropelewski, C., Wang, J., Leetma, A., Reynolds, R., Jenne, R. and Joseph, D. (1996). The NCEP/NCAR 40-year Reanalysis Project. *Bulletin of the American Meteorological Society*, **77**, 437-470,

Kalybekova, A. (2014). *Tajikistan: Power Plant Produces Heat and Anger*. Available: http://www.eurasianet.org/node/68012. Last accessed 16/10/2017.

Karaseva, M.O., Prakash, S. & Gairola, R. (2012). Validation of high-resolution TRMM-3B43 precipitation product using rain gauge measurements over Kyrgyzstan. *Theoretical and Applied Climatology*, **108**, 1-2, 147-157.

Kaser, G., Grosshauser, M. and Marzeion, B. (2010). Contribution potential of glaciers to water availability in different climate regimes. *Proceedings of the National Academy of Sciences of the United States of America* , **107**, 47, 20223–20227.

Kemmerikh, A.O. (1972). Rol lednikov v stoke rek Sredney Asii (Glaciers' part in the stream flow in Central Asia). *Data of Glaciological Studies.* **20**, 82-94 (in Russian).

Kennedy, A. M., Garen, D. C. and Koch, R. W. (2009). The association between climate teleconnection indices and Upper Klamath seasonal streamflow: Trans-Nino Index. *Hydrological Processes*, **23**, 973– 984.

Khouakhi, A. (2017). *Implementation of a source-to-sink model in R to estimate runoff from the basin upstream of Nurek reservoir, Tajikistan*. Report prepared on behalf of Acclimatise, Oxford.

Kirchhoff, C.J., Lemos, M.C. and Engle, N.L. (2013). What influences climate information use in water management? The role of boundary organizations and governance regimes in Brazil and the U.S. *Environmental Science and Policy*, **26**, 6–18.

Kistler, R., Kalnay, E., Collins, W., Saha, S., White, G., Woollen, J., Chelliah, M., Ebisuzaki, W., Kanamitsu, M., Kousky, V., van den Dool, H., Jenne, R. and Fiorino, M. (2001). The NCEP–NCAR 50-year reanalysis: Monthly means CD-Rom and documentation. *Bulletin of the American Meteorological Society*, **82**, 247–267

Kohavi, R. (1995). A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. *Proceedings of the 14th International Joint Conference on Artificial Intelligence.* 1137-1145.

Kraak, E. (2012). *Central Asia's dam debacle.* Available: http://www.chinadialogue.net/article/show/single/en/4790-Central-Asia-s-dam-debacle. [Accessed 07/12/2013].

Krause, P., Boyle, D.P. and Base, F. (2005). Comparison of different efficiency criteria for hydrological model assessment. *Advances in Geosciences*, **5**, 89-97.

Kriegel, D., Mayer, C., Hagg, W., Vorogushyn, S., Duethmann, D., Gafurov, A. and Farinotti, D. (2013). Changes in glacierisation, climate and runoff in the second half of the 20th century in the Naryn basin, Central Asia. *Global and Planetary Change*, **110**, 51-61.

Krishnaswamy, J., Vaidyanathan, S., Rajagopalan, B., Bonell, M., Sankaran, M., Bhalla, R. and Badiger, S. (2014). Non-stationary and non-linear influence of ENSO and Indian Ocean Dipole on the variability of Indian monsoon rainfall and extreme rain events, *Climate Dynamics,* **45**, 175–184.

Kumar, K.K., Rajagopalan, B. and Cane, M.A. (1999). On the weakening relationship between the Indian monsoon and ENSO. *Science,* **284**, 2156–2159.

Kure, S., Jang, S., Ohara, N., Kavvas, M.L. and Chen, Z.Q. (2013). Hydrologic impact of regional climate change for the snowfed and glacierfed river basins in the Republic of Tajikistan: hydrological response of flow to climate change. *Hydrological Processes*, **27**, 26, 4057–4070.

Kutazov, S. and Shahgedanova, M. (2009). Glacier retreat and climatic variability in the eastern Terskey–Alatoo, inner Tien Shan between the middle of the 19th century and beginning of the 21st century. *Global and Planetary Change*, **69**, 1-2, 59-70.

Laldjebaev, M. (2010). The Water-Energy Puzzle in Central Asia: The Tajikistan Perspective. *International Journal of Water Resources Development*, **26**, 23-36.

Lemos, M.C. (2008). What Influences Innovation Adoption by Water Managers? Climate Information Use in Brazil and the United States. *Journal of the American Water Resources Association*, **44**, 6, 1388-1396.

Lettenmaier, D. (2017). Observational breakthroughs lead the way to improved hydrological predictions. *Water Resources Research*, **53**, 2591-2597.

Kvasov, D.D. (1959). Causes of flow cessation in the Uzboy channel and the Aral Sea problem. *Problemy osvoyeniya pustyn,* **6**, 24-29 (in Russian).

Kvasov, D.D. and Mamedov, E.D. (1991). *The development of Irrigation and its influence on water balance.* In: Sevastyanov, D.V. History of Lakes Sevan, Issyk-Kul, Balkhash, Zaisan and Aral. Leningrad, 227-230, (in Russian).

Libertino, A., Sharma, A., Lakshmi, V. and Claps, P. (2016). A global assessment of the timing of extreme rainfall from TRMM and GPM for improving hydrologic design. *Environmental Research Letters*, **11**, 054003, doi:10.1088/1748-9326/11/5/054003.

Lorenz, C. and Kuntsmann, H. (2012). The Hydrological Cycle in Three State-of-the-Art Reanalyses: Intercomparison and Performance Analysis. *Journal of Hydrometeorology*. **13**, 1397-1420.

Ludwig, F. (2009). *Use of seasonal climate forecasts for water management.* In: Ludwig, F., Kabat, P., van Schaik, H. and van der Valk, M. Climate Change Adaptation in the Water Sector. Abingdon: Earthscan. 79-87.

Lutz, A.F., Immerzeel, W.W., Gobiet, A., Pellicciott, A. and Bierkens, M.F.P. (2013). Comparison of climate change signals in CMIP3 and CMIP5 multi-model ensembles and implications for Central Asian glaciers. *Hydrology and Earth System Sciences*, **17**, 3661–3677.

Lyon, B. and Barnston, A.G. (2005). ENSO and the Spatial Extent of Interannual Precipitation Extremes in Tropical Land Areas. *Journal of Climate*, **18**, 5095-5109.

Magar, R.B. & Jothiprakash, V. (2011). Intermittent reservoir daily-inflow prediction using lumped and distributed data multi-linear regression models. Journal of Earth System Science, **120**, 6, 1067-1084.

Mantua, N., Dettinger, M.D., Pagano, T.C., Wood, A.W., Redmond, K. and Restrepo, P. (2008). A Description *and Evaluation of Hydrologic and Climate Forecast and Data Products that Support Decision-Making for Water Resource Managers. In: Decision-Support Experiments and Evaluations using Seasonal-to-Interannual Forecasts and Observational Data: A Focus on Water Resources*. A Report by the U.S. Climate Change Science Program and the Subcommittee on Global Change Research [Beller-Simms, N., Ingram, H., Feldman, D., Mantua, N., Jacobs, K.L. and Waple, A.M. (eds.)]. NOAA's National Climatic Data Center, Asheville, NC, pp. 29-64.

Marchane, A., Jarlan, L., Boudhar A., Tramblay, Y. and Hanich, L. (2016). Linkages between snow cover, temperature and rainfall and the North Atlantic Oscillation over Morocco. *Climate Research*, **69**, 229-238.

Mariotti, A. (2007). How ENSO impacts precipitation in southwest central Asia. *Geophysical Research Letter*s, **34**, doi: 10.1029/2007GL030078.

Marshall, J., Kushnir, Y., Battisti, D., Chang, P., Czaja, A., Dickson, R., Hurrell, J., McCartney, M., Saravanan, R., Visbeck, M. (2001). North Atlantic climate variability: phenomena, impacts and mechanisms. *International Journal of climatology*, **21**, 15, 1863–1898.

Matzler, C. and Standley, A. (2000). Technical note: relief effects for passive microwave remote sensing. *International Journal of Remote Sensing*, **21**, 2403–2412.

Maurer, E.P. and Lettenmaier, D.P. (2003). Predictability of seasonal runoff in the Mississippi River basin. *Journal of Geophysical Research*, **108**, D16, doi:10.1029/2002JD002555.

McGregor, G. (2017). Hydroclimatology, modes of climatic variability and stream flow, lake and groundwater level variability. *Progress in Physical Geography*, **41**, 4, 496-512.

McMillan, H., Krueger, T. and Freer, J. (2012). Benchmarking observational uncertainties for hydrology: rainfall, river discharge and water quality. *Hydrological Processes*, **26**, 26, 4078–4111.

Mendoza, P.A., Wood, A.W., Clark, E., Rothwell, E., Clark, M.P., Nijssen, B., Brekke, L.D. and Arnold, J.R. (2017). An intercomparison of approaches for improving operational seasonal streamflow forecasts, *Hydrology and Earth System Sciences*, **21**, 3915-3935.

Menga, F. (2015). Building a nation through a dam the case of Rogun in Tajikistan. *The Journal of Nationalism and Ethnicity*, **43**, 3, 479-494.

Michel, C. (2016). *Investors Needed for Kyrgyz Hydropower Projects.* Available: http://thediplomat.com/2016/01/investors-needed-for-kyrgyz-hydropower-projects/. Last accessed 04/09/2017.

Murphy, A.H. (1993). What is a good forecast? An essay on the nature of goodness in weather forecasting. *Weather and forecasting,* **8**, 2, 281-293.

NASA. (2011). *TRMM Senior Review Proposal.* Available: http://pmm.nasa.gov/sites/default/files/document_files/TRMMSenRevProp_v1.2.pdf. Last accessed 29/05/2017.

Pagano, T.C., Wood, A.W., Ramos, M.-H., Cloke, H.L., Pappenberger, F., Clark, M.P., Cranston, M., Kavetski, D., Mathevet, T., Sorooshian, S. and Verkade. J.S. (2014). Challenges of operational river forecasting. *Journal of Hydrometerology*, **15**, 1692-1707.

Pal, I., Lall, U., Robertson, A.W., Cane, M.A. and Bansal, R. (2013). Predictability of Western Himalayan river flow: melt seasonal inflow into Bhakra Reservoir in Northern India. *Hydrology and Earth System Sciences*, **17**, 2131–2146.

Pereira-Cardenal, S.J., Riegels, N.D., Berry, P.A.M., Smith, R.D., Yakovlev, A., Siegfried, T.U. and Bauer-Gottwein, P. (2011). Real-time remote sensing driven river basin modeling using radar altimetry. *Hydrology and Earth System Sciences*, **15**, 241-254.

Petty, G.W. (2001). Physical and microwave radiative properties of precipitating clouds. Part II: A parametric 1D rain-cloud model for use in microwave radiative transfer simulations. *Journal of Applied Meteorology*, **40**, 12, 2115-2129.

Pohl, E., knoche, M., Gloaguen, R., Andermann, C. and Krause, P. (2015). Sensitivity analysis and implications for surface processes from a hydrological modelling approach in the Gunt catchment, high Pamir Mountains. *Earth surface Dynamics*, **3**, 333-362.

Radio Free Europe. (2016). *Tajikistan Introduces Electricity Rationing As Winter Nears.* Available: https://www.rferl.org/a/tajikistan-electricity-rationing-winter/28088179.html. Last accessed 16/10/2017.

Rana, S., McGregor, J. and Renwick, J. (2017). Wintertime precipitation climatology and ENSO sensitivity over central southwest Asia. *International Journal of Climatology*, **37**, 3, 1494–1509.

Rayner, S., Lach, D. and Ingram, H. (2005). Weather Forecasts are for Wimps: Why Water Resource Managers Do Not Use Climate Forecasts. *Climatic Change,* **69**, 2, 197-227.

Redmond, K.T. and Koch, R.W. (1991). Surface Climate and Streamflow Variability in the Western United states and their Relationship to Large-Scale Circulation Indices. *Water Resources Research*, **27**, 9, 2381-2399.

Robertson, D. (2016). *How good is my forecasting method? Some thoughts on forecast evaluation using cross-validation based on Australian experiences.* Available: https://hepex.irstea.fr/how-good-is-my-forecasting-method-some-thoughts-on-forecast-evaluation-using-cross-validation-based-on-australian-experiences/. Last accessed 01/09/2017.

Savoskul, O. Chevnina, E.V., Perziger, F.I., Vasilina, L.U., Baburin, V.L., Danshin A.I., Matyakubov, B. and Murakaev, R.R. (2003). *Water, climate, food, and environment in the Syr Darya Basin, in Adaptation Strategies to Changing Environments.* Institute for Environmental Studies. Vrije University., Amsterdam.

Savoskul, O.S. and Smakhtin, V (2013). *Glacier Systems and Seasonal Snow Cover in Six Major Asian River Basins: Hydrological Role under Changing Climate.* Colombo, Sri Lanka: International Water Management Institute (IWMI) Research Report 150.

Schär, C., Vasilina, L., Pertziger, F. & Dirren, S. (2004). Seasonal runoff forecasting using precipitation from meteorological data assimilation systems. *Journal of Hydrometeorology*, **5**, 959-973.

Scheel, M., Rohrer, M., Huggel, C., Santos Villar, D., Silvestre, E. & Huffman, G. (2011). Evaluation of TRMM Multi-satellite Precipitation Analysis (TMPA) performance in the Central Andes region and its dependency on spatial and temporal resolution. *Hydrology and Earth System Sciences*, **15**, 8, 2649-2663.

Schepen, A. and Wang, Q.J. (2015). Model averaging methods to merge operational statistical and dynamic seasonal streamflow forecasts in Australia. *Water Resources Research*, **51**, 1797– 1812.

Schiemann, R., Glazirina, M.G. & Schär, C. (2007). On the relationship between the Indian summer monsoon and river flow in the Aral Sea basin. *Geophysical Research Letters*, **34**, 5.

Schiemann, R., Lüthi, D., Vidale, P.L. & Schär, C. (2008). The precipitation climate of Central Asia—intercomparison of observational and numerical data sources in a remote semiarid region. *International Journal of Climatology*, **28**, 3, 295-314.

Schöne T., Dusik E., Illigner J. and Klein I. (2017). *Water in Central Asia: Reservoir Monitoring with Radar Altimetry Along the Naryn and Syr Darya Rivers.* In: International Association of Geodesy Symposia. Springer, Berlin, Heidelberg.

Schöne, T., Zech, C., Unger-Shayesteh, K., Rudenko, V., Thoss, H., Wetzel, H. U., Gafurov, A., Illigner, J., and Zubovich, A. (2013). A new permanent multi-parameter monitoring network in Central Asian high mountains - from measurements to data bases. *Geoscientific Instrumentation, Methods and Data Systems*, **55**, 2, 97-111.

Sehring, J. and Diebold, A. (2012). From Glaciers to the Aral Sea – water unites. Trescher Verlag GmbH, Berlin.

Sene, K. (2016). *Hydrometeorology- Forecasting and Applications*. 2nd ed. Switzerland: Springer International Publishing.

Sevruk, B., Ondras, M. and Chvila, B. (2009). The WMO precipitation measurement intercomparisons. *Atmospheric Research*, **92**, 3, 376–380.

Siegfried, T. Bernauer, T., Guiennet, R., Sellars, S., Robertson, A.W., Mankin, J., Bauer-Gottwein, P. and Yakovlev, A. (2012). Will climate change exacerbate water stress in Central Asia? *Climatic Change*, **112**, 881–899.

Singh, P. and Bengtsson, L. (2005). Impact of warmer climate on melt and evaporation for the rainfed, snowfed and glacierfed basins in the Himalayan region. *Journal of Hydrology*, **300**, 140–154.

Sommer, R., Glazirina, M., Yuldashev, T., Otarov, A., Ibraeva, M., Martynova, L., Bekenov, M., Kholov, B., Ibragimov, N. & Kobilov, R. (2013). Impact of climate change on wheat productivity in Central Asia. *Agriculture, Ecosystems & Environment*, **178**, 78-99.

Sorg, A., Bolch, T., Stoffel, M., Solomina, O. & Beniston, M. (2012). Climate change impacts on glaciers and runoff in Tien Shan (Central Asia). *Nature Climate Change*, **2**, 10, 725-731.

Sorg, A., Huss, M., Rohrer, M. and Stoffel, M. (2014b). The days of plenty might soon be over in glacierized Central Asian catchments. *Environmental Research Letters*, **9**, doi:10.1088/1748-9326/9/10/104018.

Sorg, A., Mosello, C., ShalpykovaG., Allan, A., Hill Clarvis, M. and Stoffel, M. (2014). Coping with changing water resources: The case of the Syr Darya river basin in Central Asia. Environmental Science and Policy, **43**, 68-77.

Su, F., Hong, Y. and Lettenmaier, D.P. (2008). Evaluation of TRMM Multisatellite Precipitation Analysis (TMPA) and Its Utility in Hydrologic Prediction in the La Plata Basin. *Journal of Hydrometeorology*, **9**, 622-640.

Svensson, C. (2016). Seasonal river flow forecasts for the United Kingdom using persistence and historical analogues. *Hydrological Sciences journal*, **61**, 1, 19-35.

Svensson, C., Brookshaw, A., Scaife, A. A., Bell, V.A., Mackay, J.D., Jackson, C.R., Hannahford, J., Davies, H.N., Arribas, A. and Stanley, S. (2015). Long-range forecasts of UK winter hydrology. *Environmental Research Letters*, **10**, 064006, doi:10.1088/1748-9326/10/6/064006.

Syed, F., Giorgi, F., Pal, J. & King, M. (2006). Effect of remote forcings on the winter precipitation of central southwest Asia part 1: observations. *Theoretical and applied climatology*, **86**, 1-4, 147-160.

Syed, F.S., Giorgi, F., Pal, J.S. and Keay, K. (2010). Regional climate model simulation of winter climate over Central–Southwest Asia, with emphasis on NAO and ENSO effects. *International Journal of Climatology*, **30**, 220–235.

Teegavarapu, R. & Elshorbagy, A. (2005). Fuzzy set based error measure for hydrologic model evaluation. *Journal of Hydroinformatics*, **7**, 199-208.

Tetzlaff, D., Carey, S. K., McNamara, J. P., Laudon, H. and Soulsby, C. (2017). The essential value of long-term experimental data for hydrology and water management. *Water Resources Research*, **53**, 2598–2604.

Thornes, J.E. and Stephenson, D.B. (2001). How to judge the quality and value of weather forecast products. *Meteorological Applications,* **8**, 307–314.

Tippett, M., Goddard, L. and Barnston, A.G. (2005). Statistical–Dynamical Seasonal Forecasts of Central-Southwest Asian Winter Precipitation. *Journal of Climate*, **18**, 1831-1843.

Tolstov, S.P. (1962)*. Through the Ancient Deltas of Oxus and Jaxartes*. Moscow, 1-324 (in Russian).

UNDP. (2012). *Natural Disaster Risks in Central Asia*. Available: http://europeandcis.undp.org/uploads/public1/files/vulnerability/Senior%20Economist%20Web%20site/Policy%20brief_Natural%20Disaster%20Risks%20in%20Central%20Asia.pdf. Last accessed 18/07/2013.

Unger-Shayesteh, K., Vorogushyn, S., Farinotti, D., Gafurov, A., Duethmann, D., Mandychev, A., and Merz, B. (2013) What do we know about past changes in the water cycle of Central Asian headwaters? A review, *Global and Planetary Change*, **110**, 4-25.

Unicef. (2008). *Maternal and Child Survival in Extreme Winter Conditions in Tajikistan.* Last accessed via Google on 16/10/2017.

Vila, D.A., de Goncalves, L.G.G., Toll, D. and Rozante, J.R. (2009). Statistical Evaluation of Combined Daily Gauge Observations and Rainfall Satellite Estimates over Continental South America. *Journal of Hydrometeorology*, **10**, 533-543.

Wagener, T., Wheater, S.H. and Gupta, V.K. (2004). Rainfall-runoff modelling-a review, In: *Rainfall-Runoff Modelling in Gauged and Ungauged Catchments*. London: Imperial College Press. 1-8.

Webster, P.J., Magana, V.O., Palmer, T., Shukla, J., Tomas, R., Yanai, M.u. & Yasunari, T. (1998). *Monsoons: Processes, predictability, and the prospects for prediction. Journal of Geophysical Research: Oceans (1978–2012),* **103**, 14451-14510.

Wegerich, K. (2008). *Hydro-hegemony in the Amu Darya Basin.* Water Policy, **10**, S2, 71-88.

Werner, K., Averyt, K., and Owen, G. (2013). River forecast application for water management: Oil and water? *Weather, Climate and Society*, **5**, 3, 244–253.

Werner, K., Brandon, D., Clark, M. and Gangopadhyay, S. (2004). Climate Index Weighting Schemes for NWS ESP-Based Seasonal Volume Forecasts. *Journal of hydrometeorology.* **5**, 1076-1090.

Wilby, R.L. and Dixon, S.G. (2017). Annex I – Step-by-step guide to seasonal forecasting of inflows for Nurek reservoir. In: Connell, R. and Wilby, R.L. 2017. *Tajikistan: Capacity building to strengthen the climate resilience of energy sector assets and operations. Mission Report for Training Session T1.3: Seasonal forecasting of regional hydroclimate and reservoir inflows*, August 2017. Prepared for European Bank for Reconstruction and Development.

Wilby, R.L., Clifford, N.J., Luca, P.D., Harrigan, S., Hillier, J.K., Hodgkins, R., Johnson, M.F., Matthews, T.K.R., Murphy, C., Noone, S.J., Parry, S., Prudhomme, C., Rice, S.P., Slater, L.J., Smith, K.A. and Wood, P.J. (2017). The 'dirty dozen' of freshwater science: detecting then reconciling hydrological data biases and errors. *WIREs Water*, **4**.

Wilby, R.L., Friedhoff, M., Rabb, B., Connell, R., Minikulov, N., Homidov, A., Shodmanov, M. and Leonidova, N. (2011a). *Tajikistan Pilot Programme for Climate Resilience (PPCR) Project A4 - Improving the Climate Resilience of Tajikistan's Hydropower Sector (Final report).* Acclimatise and Sinclair Knight Mertz, Oxford.

Wilby, R.L., Minikulov, N. and Rabb, B. (2011b). *ANNEX 1: Modelling the climate-sensitivity of rivers entering the Kairakkum and Nurek reservoirs, Tajikistan.* Available: http://www.ppcr.tj/IP/Phase1/Component4/Annex%201.pdf. Last accessed 31/07/2014.

Wilks, D. (2011). *Statistical Methods in the Atmospheric Sciences.* 3rd ed. USA: Elsevier. 215.

Wilson, L.J. (2004). Introduction to forecast verification [Powerpoint slides]. Available: https://www.isse.ucar.edu/ams/present/wilson.ppt. Last accessed 02/09/2017.

Wood, A.W. and Lettenmaier, D.P. (2006). A test bed for new seasonal hydrologic forecasting approaches in the western United States. *Bulletin of the American Meteorological Society*, **87**, 1699-1712.

World Bank (2012). *The Role of Hydrometeorological Services in Disaster Risk Management.* Washington D.C. 43-47. Available: http://www.unisdr.org/files/27645_webrestheroleofhydromet.pdf. Last accessed 13/10/2017.

World Bank. (2008). *Kyrgyz Republic Energy Emergency Assistance Project.* Available: http://documents.worldbank.org/curated/en/278511468266455064/pdf/E19850P101392 01BLIC10EMP00902602008.pdf. Last accessed 16/10/2017.

World Bank. (2010). *Tajikistan Energy Loss Reduction Project: Request for Inspection.* Available: http://siteresources.worldbank.org/EXTINSPECTIONPANEL/Resources/Map_taj38165. pdf. Last accessed 07/07/2017.

World Bank. (2013a). *Tajikistan's Winter Energy Crisis: electricity supply and demand alternatives.* Available: http://documents.worldbank.org/curated/en/500811468116363418/pdf/796160PUB0RE PL00Box377374B00PUBLIC0.pdf. Last accessed 15/10/2017.

World Bank. (2013b). *The Kyrgyz Republic: Poverty Update.* Available: Available: http://www-wds.worldbank.org/external/default/WDSContentServer/WDSP/IB/2013/06/26/0003330 37_201306261226558/Rendered/PDF/782130ESW0P101000Final0250June00ENG.pdf. Last accessed 07/31/2013.

World Bank. (2015). *World Bank Supports Energy Sector Development in the Kyrgyz Republic.* Available: http://www.worldbank.org/en/news/press-release/2015/01/29/world-bank-supports-energy-sector-development-in-the-kyrgyz-republic. Last accessed 16/10/2017.

World Bank. (2017). *Central Asia Hydrometeorology Modernization Project.* Available: http://projects.worldbank.org/P120788/central-asia-hydrometeorology-modernization-project?lang=en. Last accessed 27/06/2017.

World Meteorological Organisation (WMO). (2000). *Standardised Verification System (SVS) for Long-Range Forecasts (LRF).* Available: https://www.wmo.int/pages/prog/www/DPS/SVS-for-LRF.html#2.1   Long-Range Forecasts. Last accessed 2017/06/27.

Yadav, R.,K., Ramu, D.A. and Dimri, A.P. (2013). On the relationship between ENSO patterns and winter precipitation over North and Central India. *Global and Planetary Change*, **107**, 50–58.

Yadav, R.K., Rupa Kumar, K. and Rajeevan, M. (2009). Increasing influence of ENSO and decreasing influence of AO/NAO in the recent decades over northwest India winter precipitation. *Journal of Geophysical Research*, **114**, D12112, DOI:10.1029/2008JD011318.

Yadav, R.K., Yoo, J.H., Kusharski, F. and Abid, M.A. (2010). Why is ENSO influencing northwest India winter precipitation in recent decades? *Journal of Climate*, **23**, 1979-1993.

Yao, H. and Georgakakos, A. (2001). Assessment of Folsom Lake response to historical and potential future climate scenarios: 2. Reservoir management. *Journal of Hydrology*, **249**, 1-4, 176–196.

Yatagai, A., Kamiguchi, K., Arakawa, O., Hamada, A., Yasutomi, N. and 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.

Yin, Z.Y., Wang, H. and Liu, X. (2014). A Comparative Study on Precipitation Climatology and Interannual Variability in the Lower Midlatitude East Asia and Central Asia. *Journal of Climate*, **27**, 7830-7848.

Yong, B. Hong, Y., Ren, L., Gourley, J., Huffman, G.J., Chen, X., Wang, W. and Khan, S.I. (2012). Assessment of evolving TRMM-based multisatellite real-time precipitation estimation methods and their impacts on hydrologic prediction in a high latitude basin. *Journal of Geophysical Research*, **117**, 1-21.

Zaitchik, B.F., Rodell, M. and Olivera, F. (2010). Evaluation of the Global Land Data Assimilation System using global river discharge data and a source-to-sink routing scheme. *Water Resources Research*, **46**, 6, DOI: 10.1029/2009WR007811.

Zhang, H., Ouyang, Z., Zheng, H. and Wang, X. (2009). Recent climate trends on the Northern slopes of the Tianshan Mountains, Xinjiang, China. *Journal of Mountain Science*, **6**, 255–265.

Zhao, C., Yoa, S., Zhang, S., Han, H., Zhao, Q. and Yi, S. (2015). Validation of the Accuracy of Different Precipitation Datasets over Tianshan Mountainous Area. *Advances in Meteorology*, doi: 10.1155/2015/617382.

Zhou, H., Aizen, E. and Aizen, V. (2017). Constructing a long-term monthly climate data set in central Asia. *International Journal of Climatology.* doi: 10.1002/joc.5259.
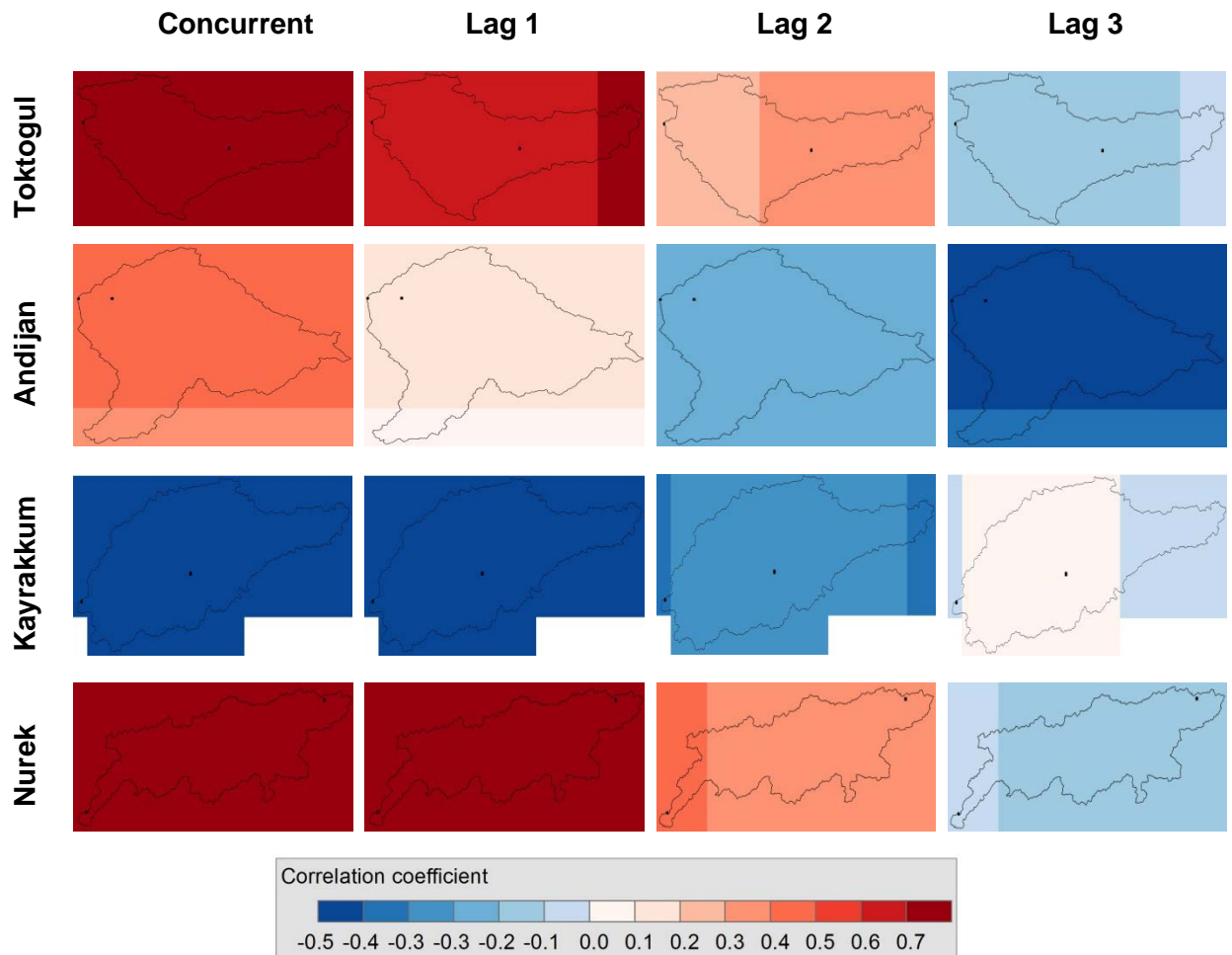
Appendix A

Figure Apx 1: Monthly observed inflow into Toktogul, Andijan, Kayrakkum and Nurek reservoirs correlated with monthly NCEP derived teperature for the years 2001-2010 for concurrent (left) and precipitation leading (right) discharge by 1, 2, and 3 months. Note that a 4 month lead was also undertaken but not shown for brevity.
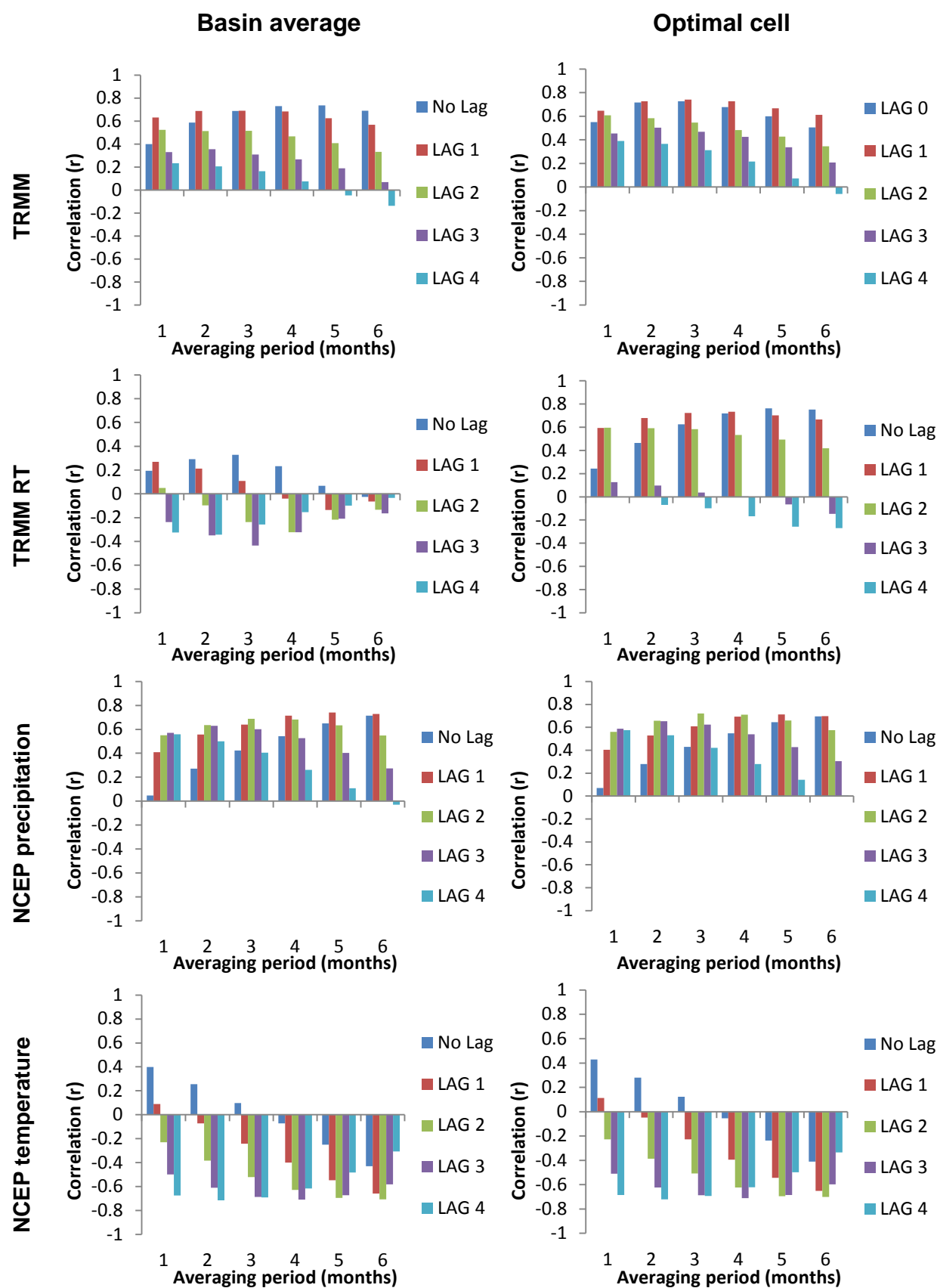
Figure Apx 2: Correlation (*r*) of gauged inflows to Andijan with lagged predictors averaged over 1-6 months. Accounting for autocorrelation at *p* = 0.05, *r*_{crit} = 0.49.

Figure Apx 3: Correlation (*r*) of gauged inflows to Kayrakkum with lagged predictors averaged over 1-6 months. Accounting for autocorrelation at *p* = 0.05, *r*_{crit} = 0.47.

Figure Apx 4: Correlation (*r*) of gauged inflows to Nurek with lagged predictors averaged over 1-6 months. Accounting for autocorrelation at *p* = 0.05, *r*_{crit} = 0.54.
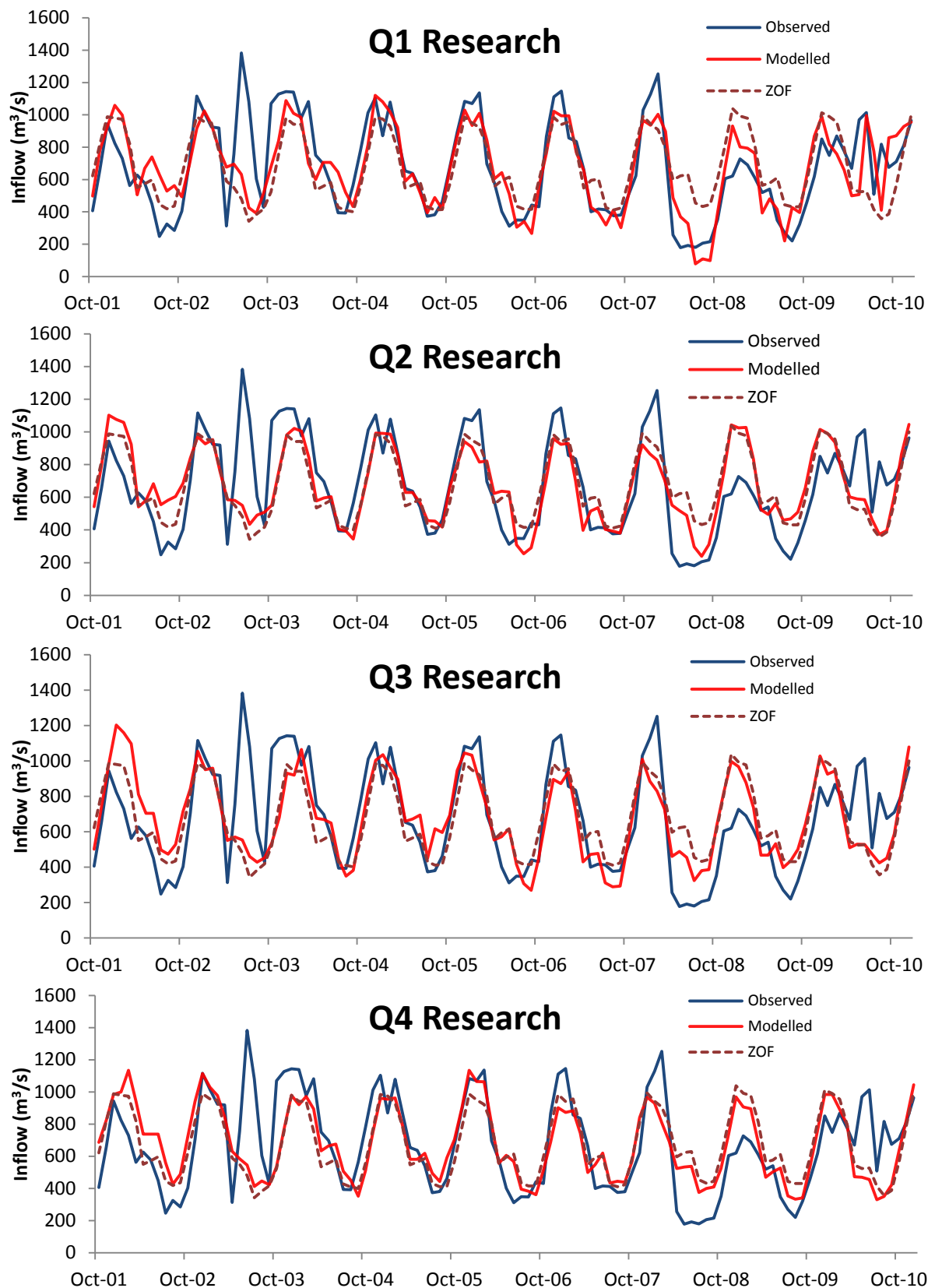
172

Figure Apx 5: Cross validated research model inflow forecasts for Toktogul with lead time one (Q1), two (Q2), three (Q3) and four (Q4) months compared to the ZOF (long term monthly mean discharge).

Figure Apx 6: Cross validated research model inflow forecasts for Toktogul with lead time one (Q1), two (Q2), three (Q3) and four (Q4) months compared to the ZOF (long term monthly mean discharge).
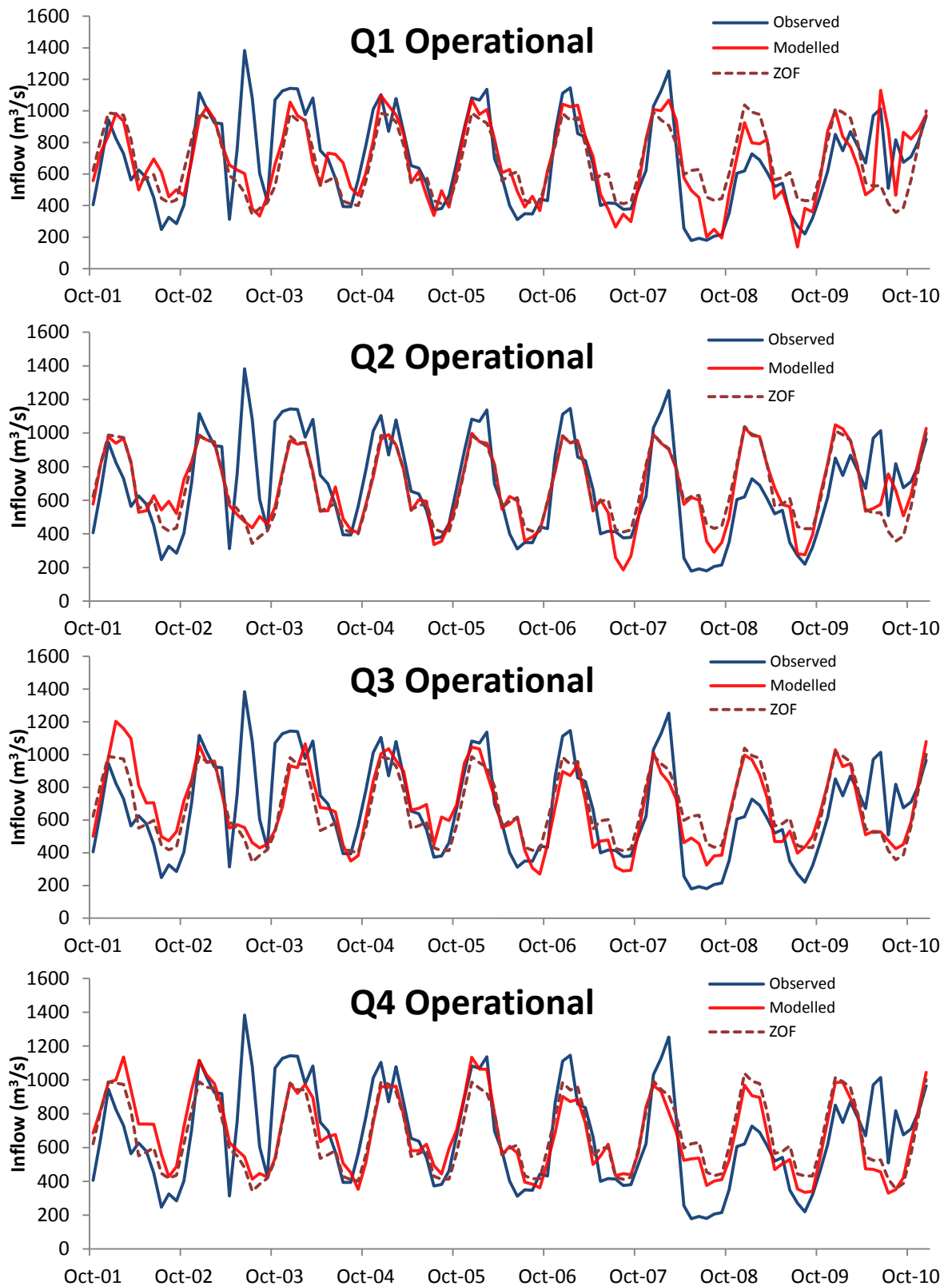
Figure Apx 7: Cross validated research model inflow forecasts for Andijan with lead time one (Q1), two (Q2), three (Q3) and four (Q4) months compared to the ZOF (long term monthly mean discharge).

Figure Apx 8: Cross validated operational model inflow forecasts for Andijan with lead time one (Q1), two (Q2), three (Q3) and four (Q4) months compared to the ZOF (long term monthly mean discharge).

Figure Apx 9: Cross validated research model inflow forecasts for Kayrakkum with lead time one (Q1), two (Q2), three (Q3) and four (Q4) months compared to the ZOF (long term monthly mean discharge).
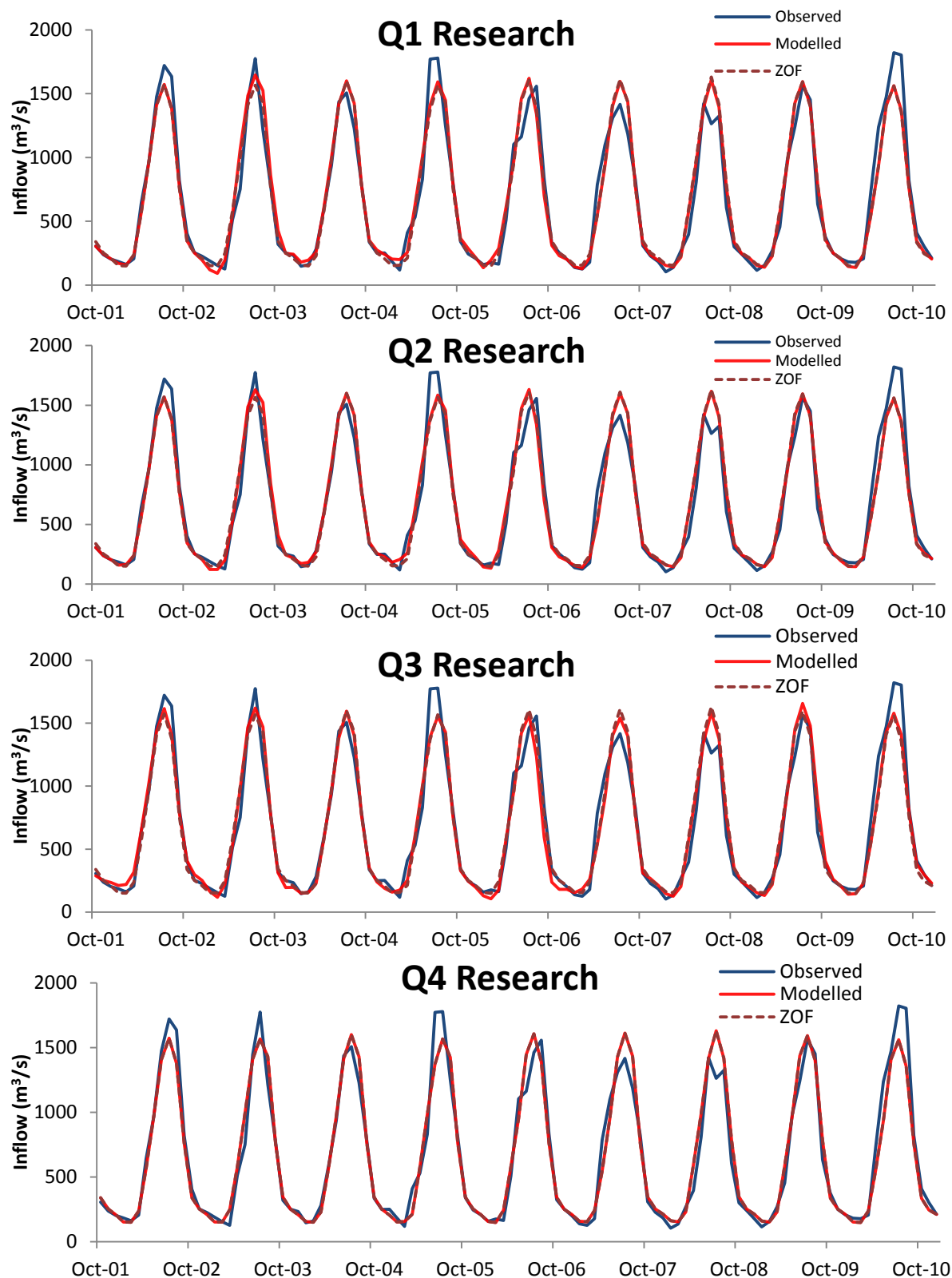
Figure Apx 10: Cross validated operational model inflow forecasts for Kayrakkum with lead time one (Q1), two (Q2), three (Q3) and four (Q4) months compared to the ZOF (long term monthly mean discharge).

Figure Apx 11: Cross validated research model inflow forecasts for Nurek with lead time one (Q1), two (Q2), three (Q3) and four (Q4) months compared to the ZOF (long term monthly mean discharge).
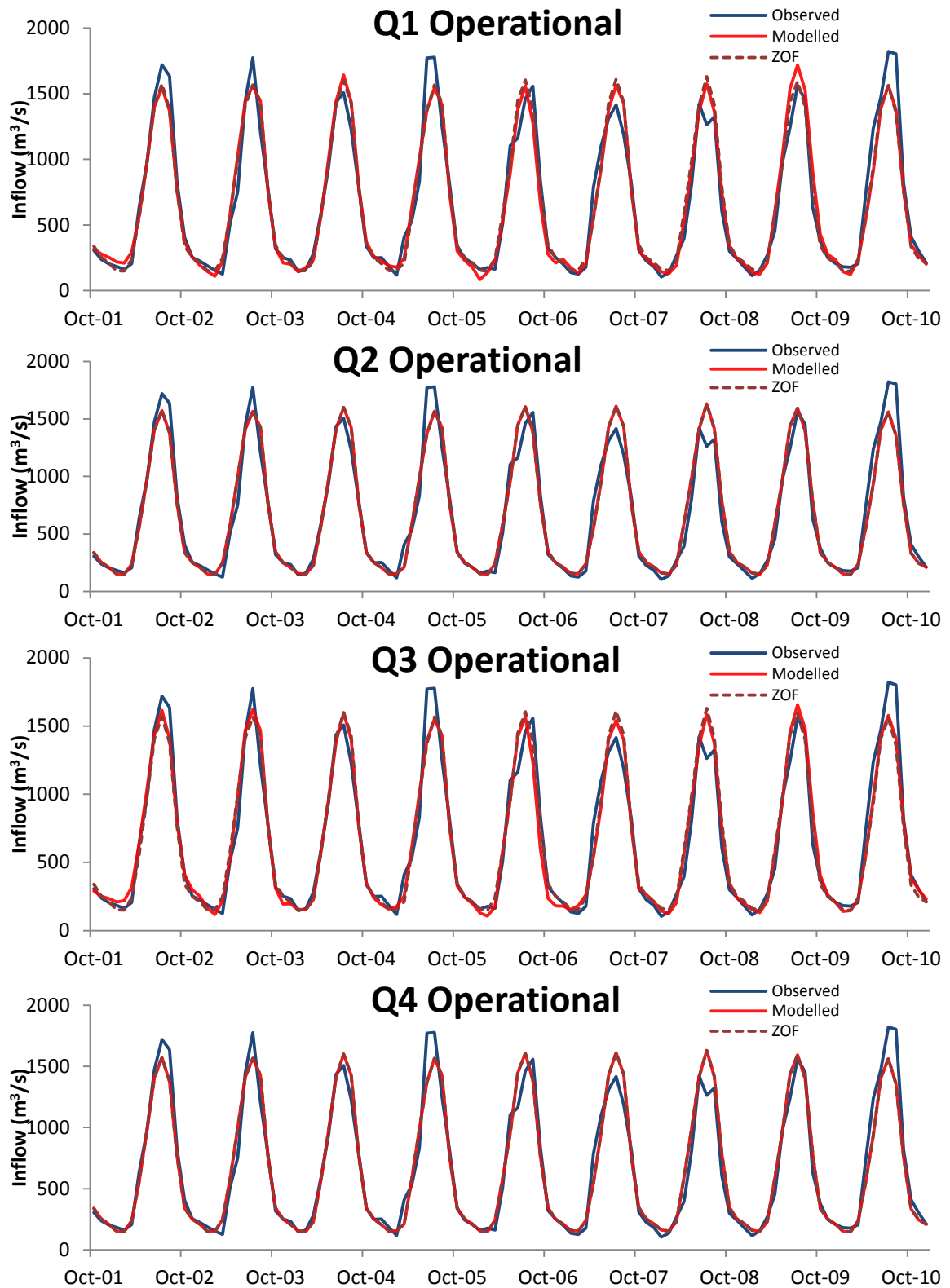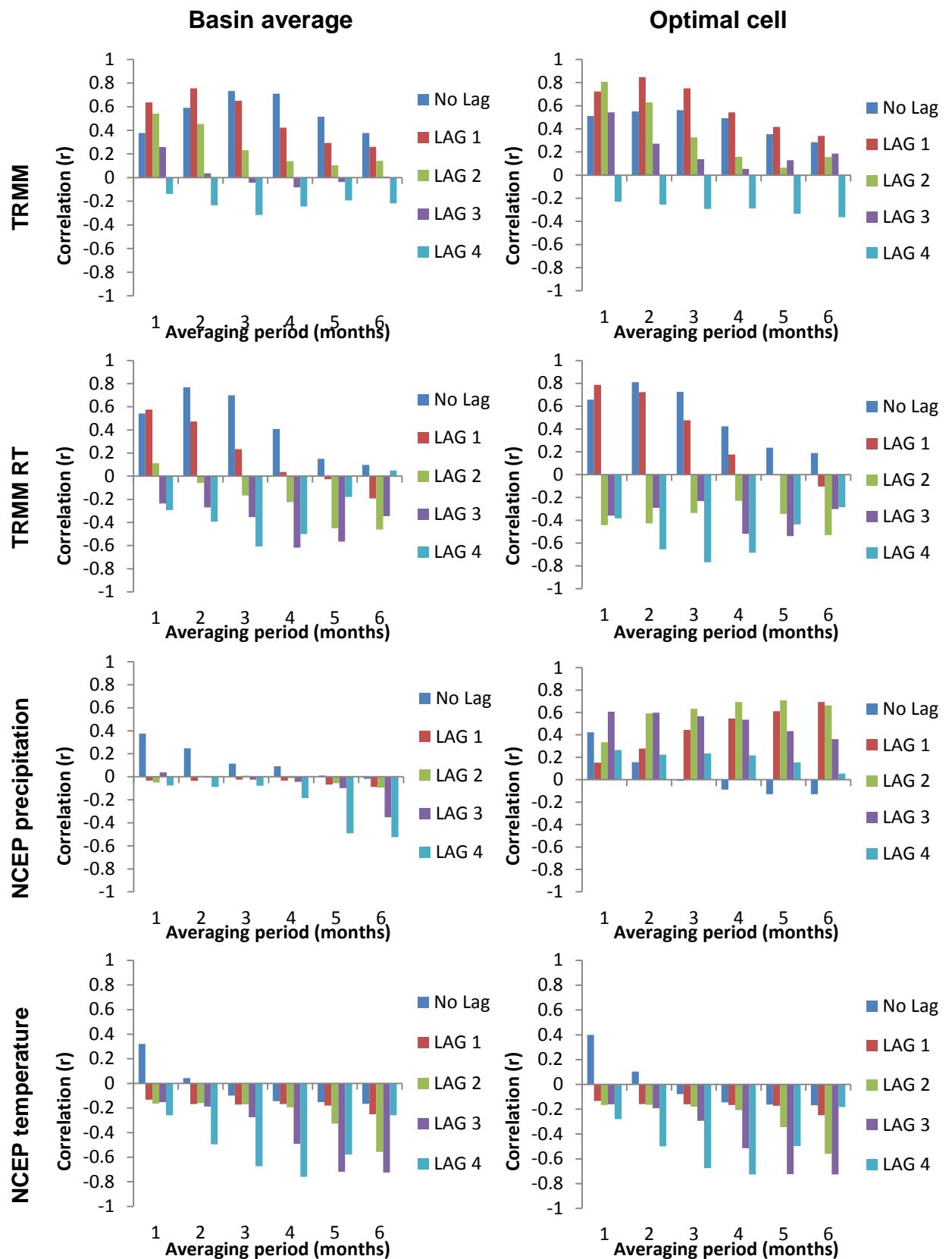
Figure Apx 12: Cross validated operational model inflow forecasts for Nurek with lead time one (Q1), two (Q2), three (Q3) and four (Q4) months compared to the ZOF (long term monthly mean discharge).

Figure Apx 13: Correlation (*r*) of gauged inflows to Toktogul (months April-September only) with lagged predictors averaged over 1-6 months for the years 2002-2010.
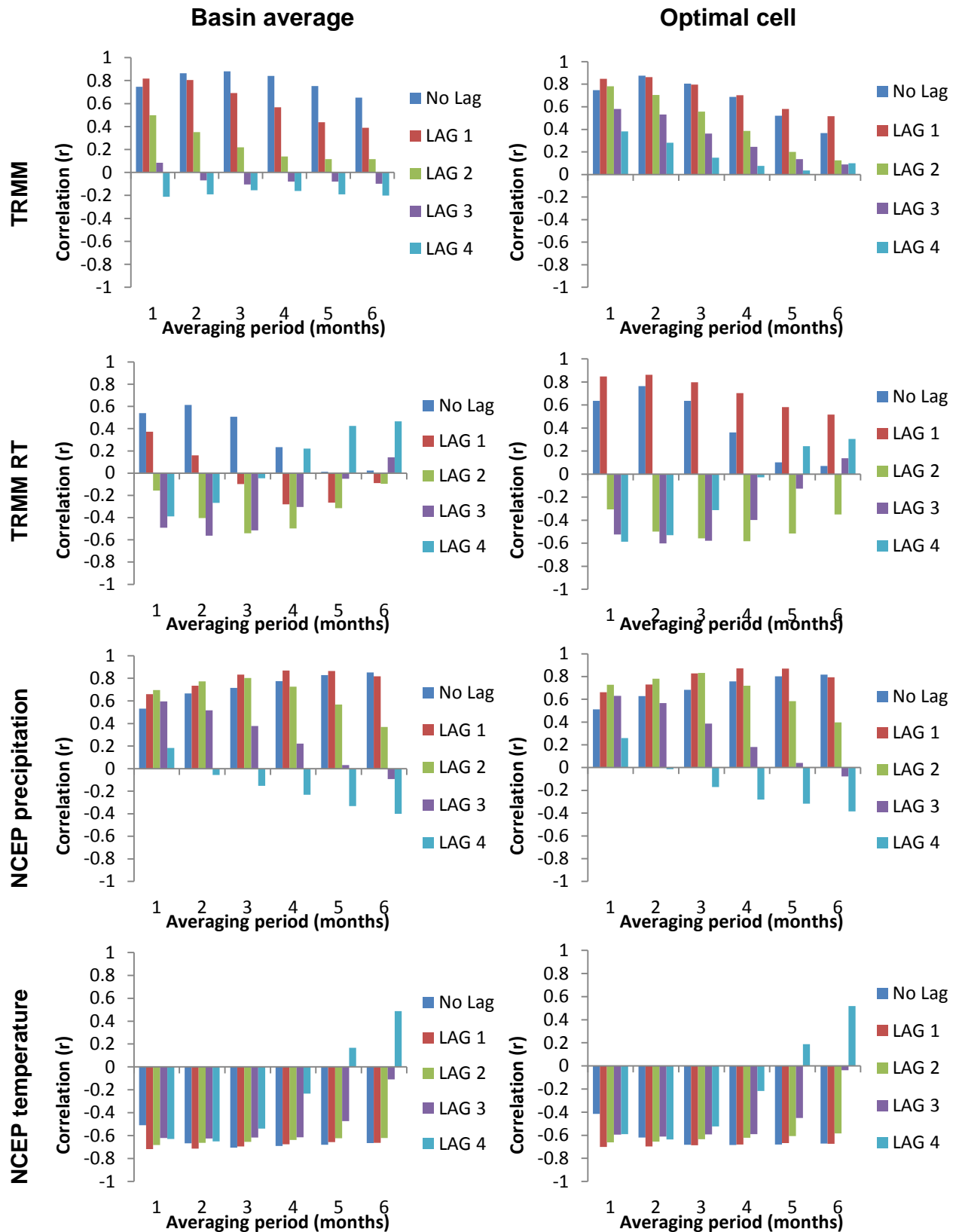
Figure Apx 14: Correlation (*r*) of gauged inflows to Andijan (months April-September only) with lagged predictors averaged over 1-6 months for the years 2002-2010.
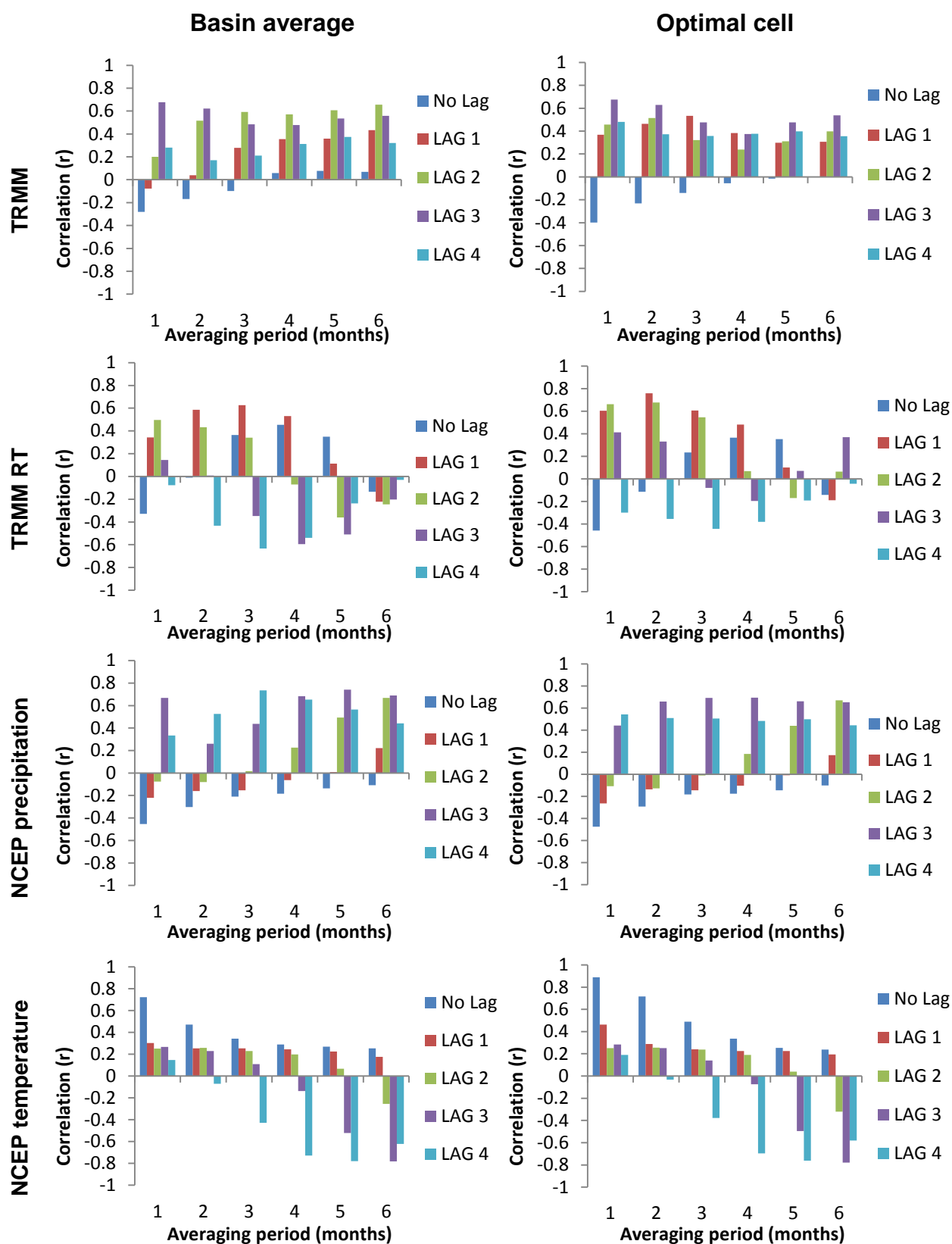
Figure Apx 15: Correlation (*r*) of gauged inflows to Nurek (months April-September only) with lagged predictors averaged over 1-6 months for the years 2002-2010.

Figure Apx 16: Predictors used stratified multiple regression models for each basin at each lead time (1-4 months) (n = 54). Notations can be found in Table 4.1.

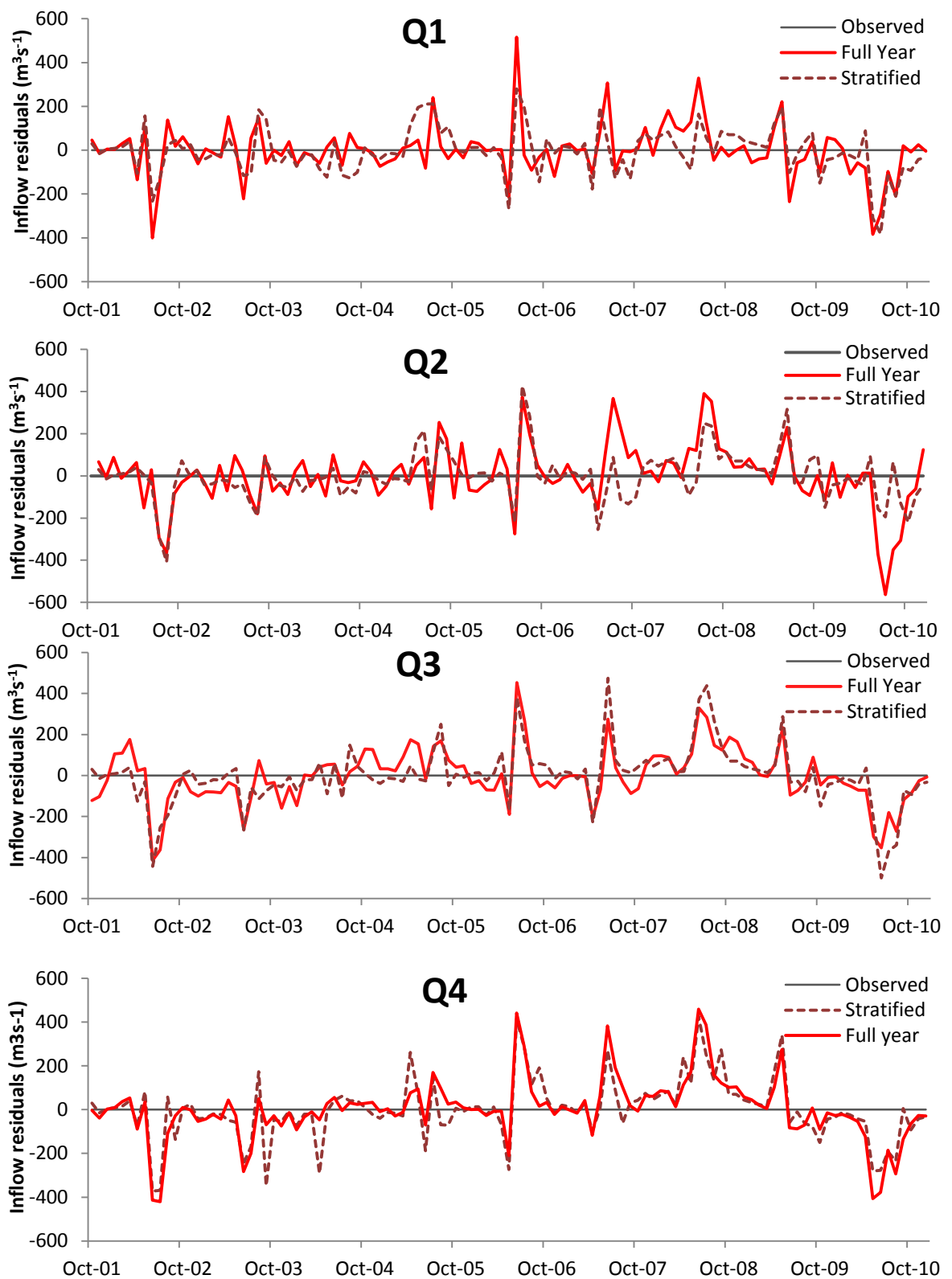|	Q1	Q2	Q3	Q4
Toktogul	M	M	M	M
	$Flow_{1,1}$	$NCEP\_P\_Op_{2,5}$	$NCEP\_P\_Op_{3,1}$	$NCEP\_T\_Op_{4,4}$
	$NCEP\_P\_Op_{1,6}$	$NCEP\_P\_Av_{2,6}$		$NCEP\_P\_Op_{4,1}$
	$TRMMRT\_Op_{1,1}$			$NCEP\_T\_Av_{4,4}$
	$NCEP\_P\_Av_{1,6}$			
Andijan	M	M	M	M
	$Flow_{1,1}$	$Flow_{2,1}$	$NCEP\_P\_Av_{3,1}$	$NCEP\_P\_Av_{4,6}$
	$NCEP\_P\_Av_{1,4}$	$NCEP\_P\_Av_{2,3}$		$NCEP\_T\_Op_{4,2}$
Nurek	M	M	M	M
	$NCEP\_P\_Av_{1,6}$		$NCEP\_P\_Op_{3,4}$	

Figure Apx 17: Cross validated operational inflow forecast residuals for Toktogul with lead time one-four months. Full year models forecast all 12 months, whereas stratified models forecast only summer inflows (April-September), using ZOF as winter flows.
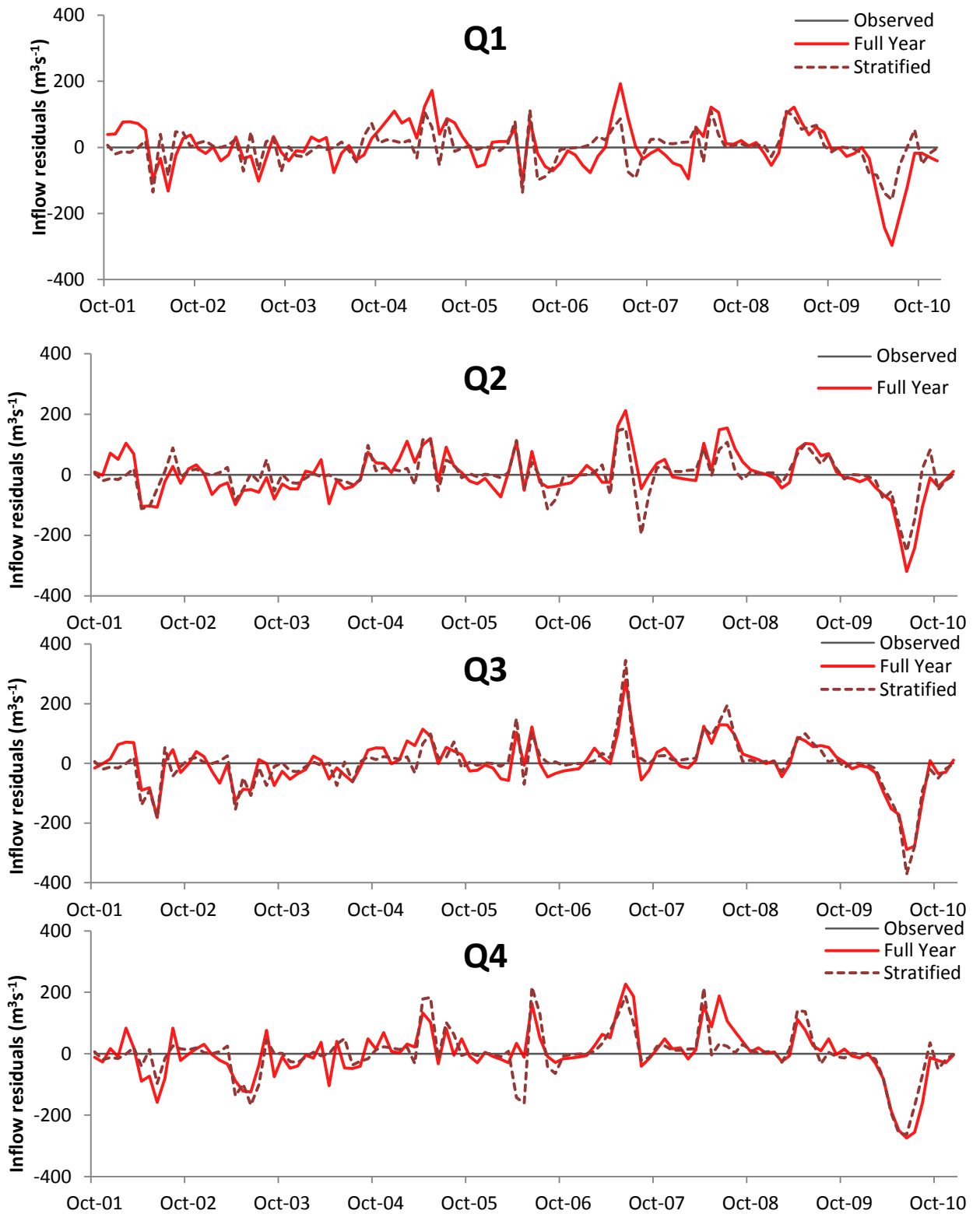
Figure Apx 18: Cross validated operational inflow forecast residuals for Andijan with lead time one-four months. Full year models forecast all 12 months, whereas stratified models forecast only summer inflows (April-September), using ZOF as winter flows.
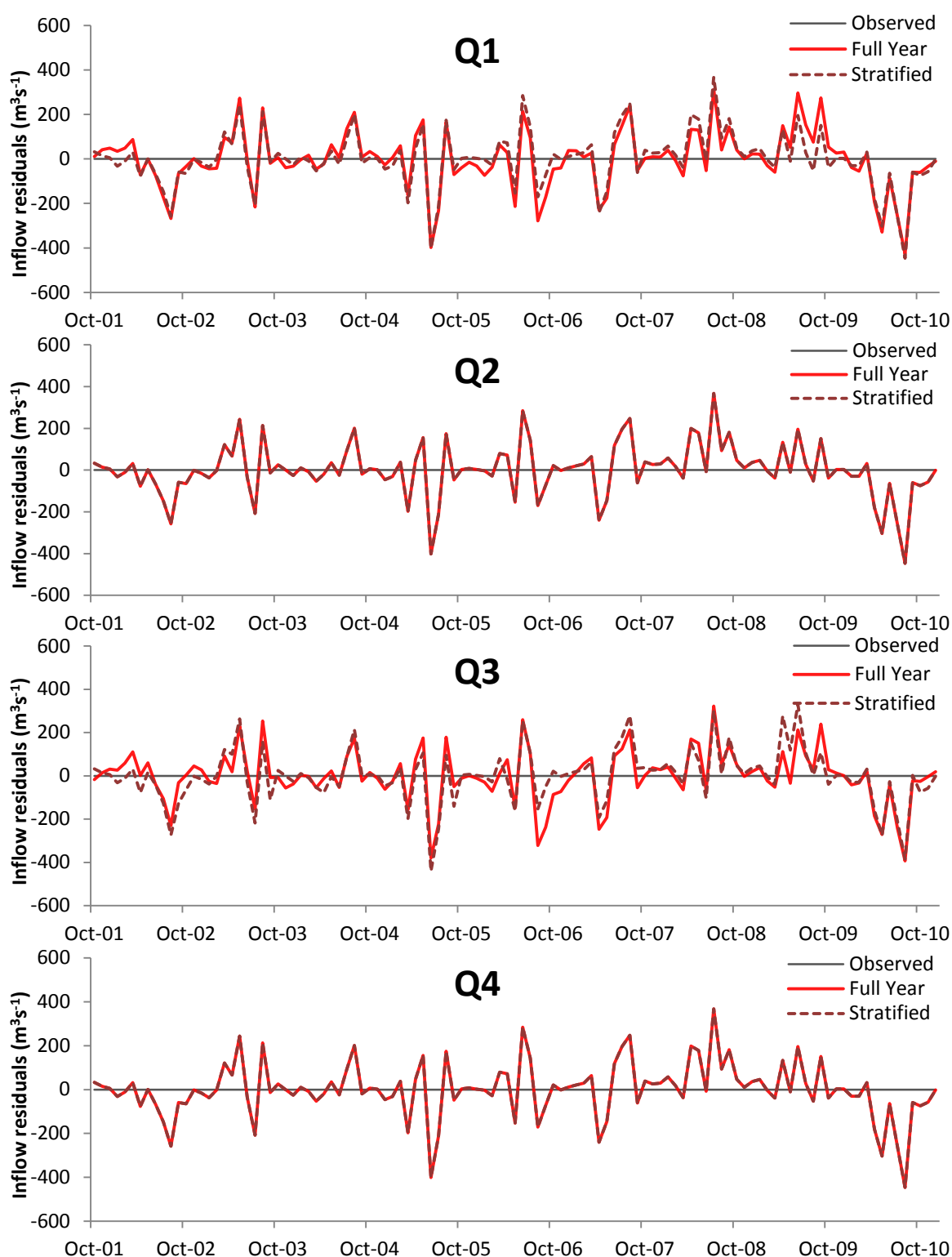
Figure Apx 19: Cross validated operational inflow forecast residuals for Nurek with lead time one-four months. Full year models forecast all 12 months, whereas stratified models forecast only summer inflows (April-September), using ZOF as winter flows.
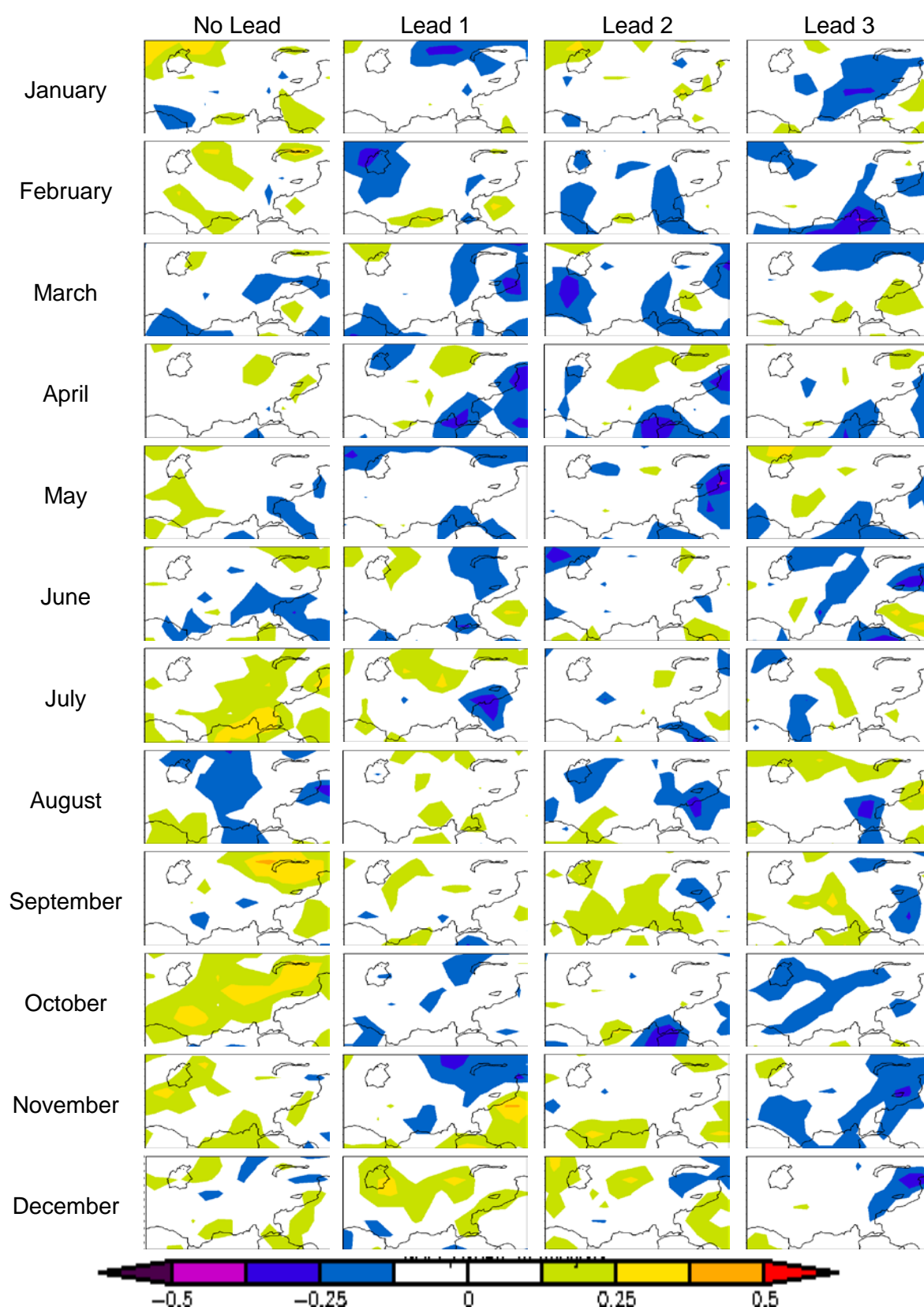
Figure Apx 20: Monthly NAO (PC) correlated with monthly NCEP precipitation rate for the years 1950-2014 at lead times 0-3 months (left column shows target precipitation month, with index month varied to account for lead time). Yellow/dark blue areas are significant at p=0.05, orange/pink are significant at p=0.01.
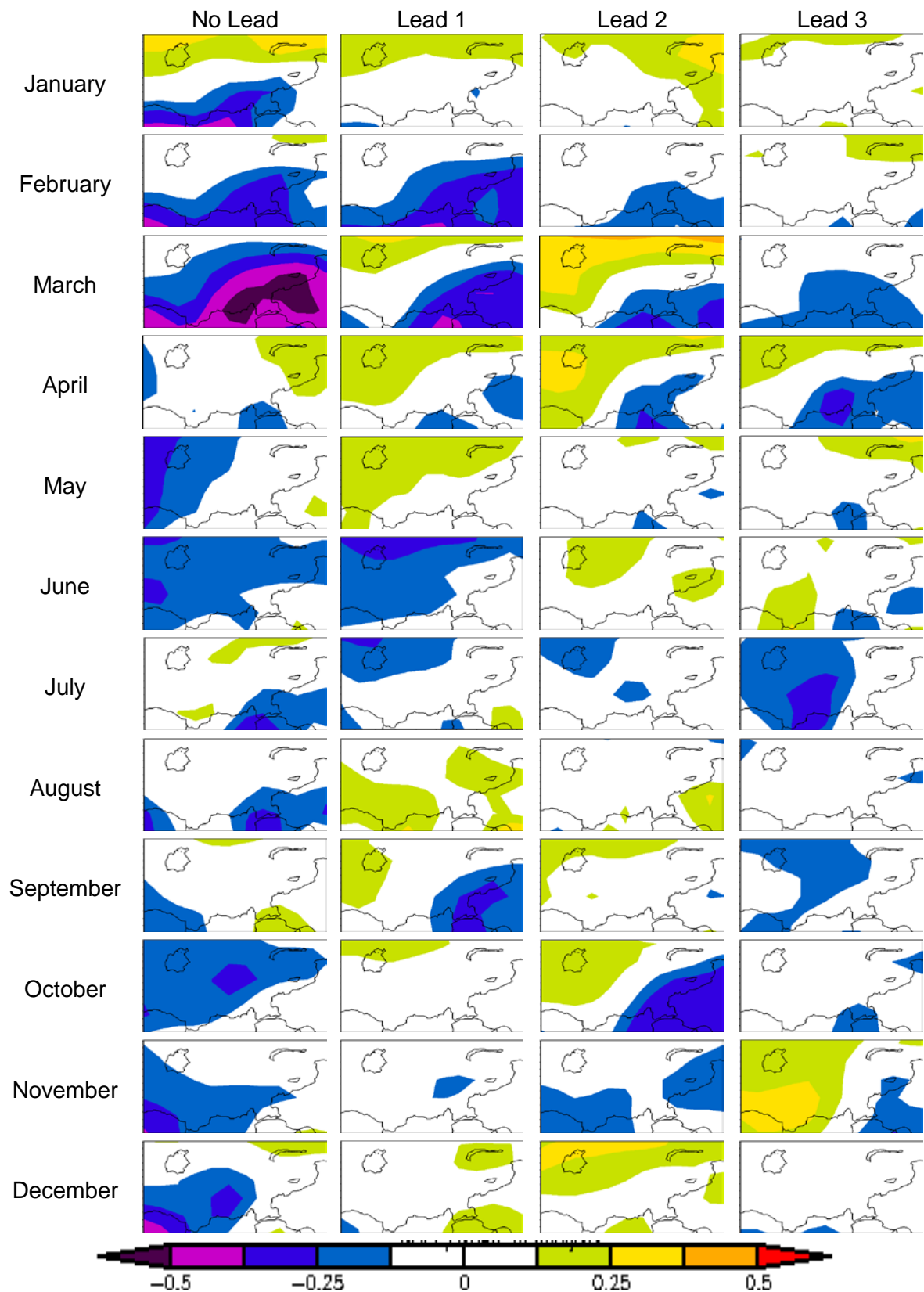
Figure Apx 21: Monthly NAO (PC) correlated with monthly NCEP temperature for the years 1950-2014 at lead times 0-3 months (left column shows target precipitation month, with index month varied to account for lead time). Yellow/dark blue areas are significant at *p*=0.05, orange/pink are significant at *p*=0.01.
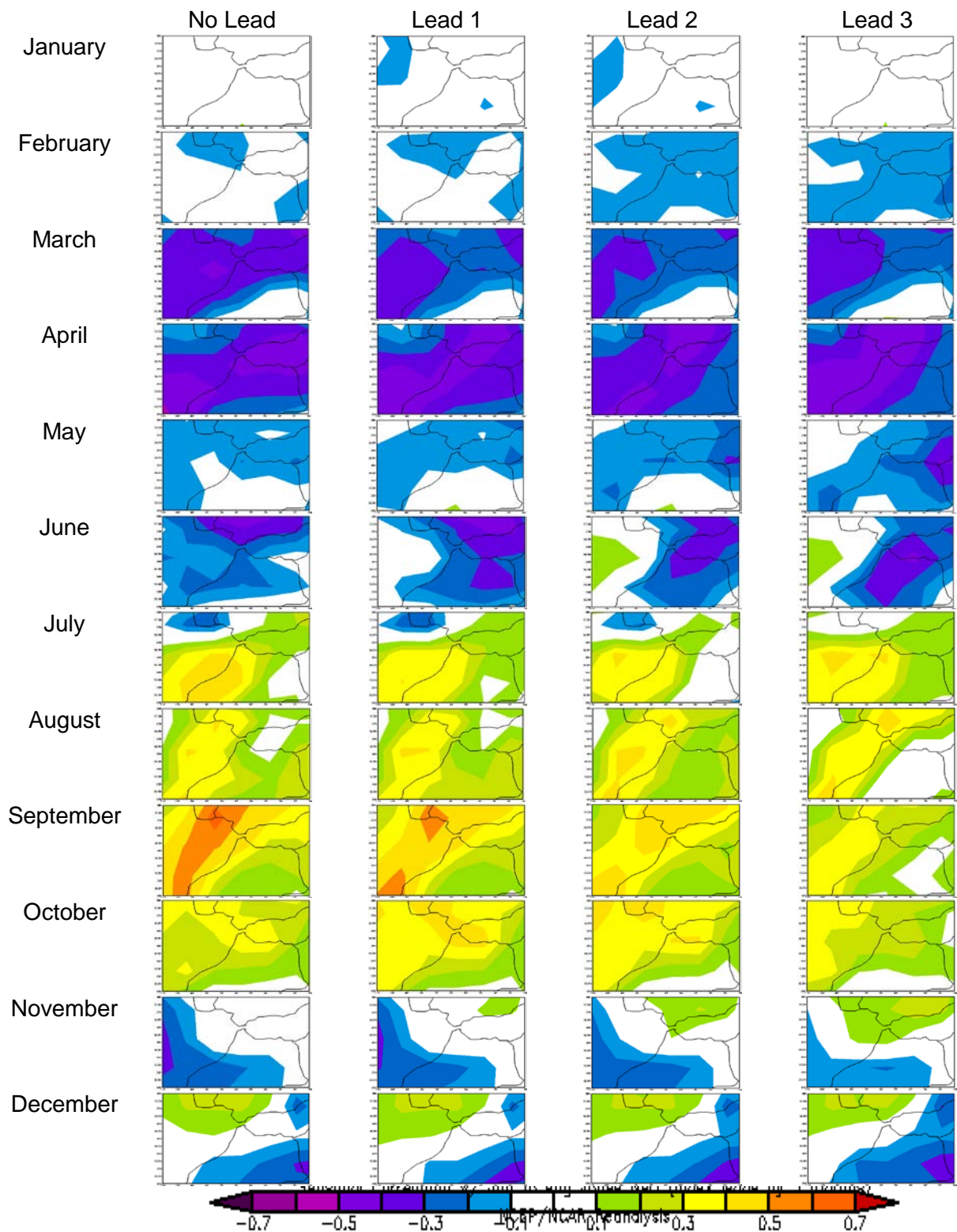
Figure Apx 22: Monthly Niño 3.4 correlated with monthly NCEP precipitation rate over Morocco for the period 1952-1989 at lead times 0-3 months (left column shows target precipitation month, with index month varied to account for lead time). Purple/Yellow areas and greater are significant at *p*=0.05.

Appendix B - Step-by-step guide to seasonal forecasting of inflows for Nurek reservoir (Wilby and Dixon, 2017).
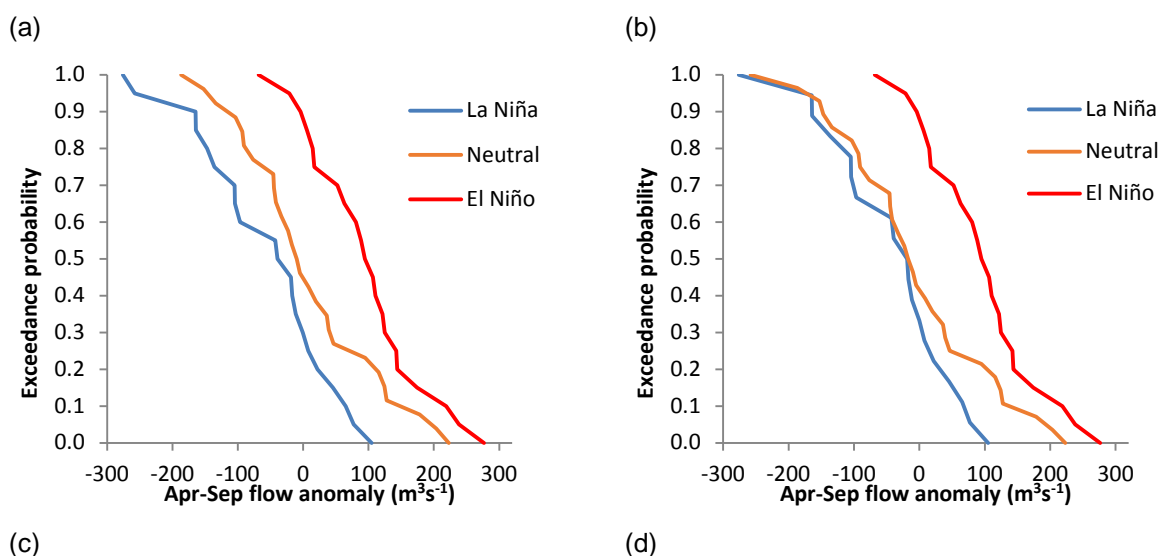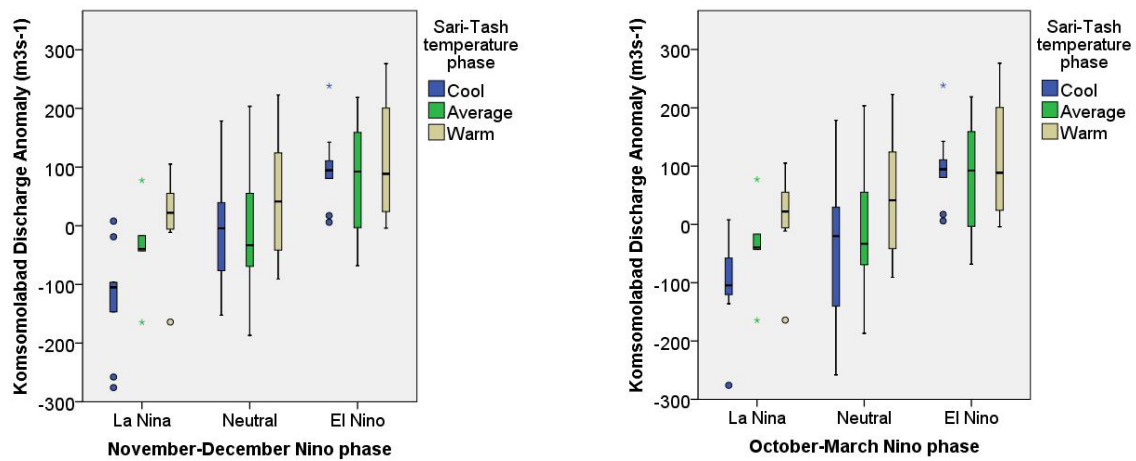
Technical Annex to Workshop Report

Seasonal forecasts of inflows for Nurek reservoir

This Technical Annex provides some background information then step-by-step guidance on how to produce a seasonal forecast of mean summer inflows anomalies to Nurek. The statistical model is described by Dixon and Wilby (2017).

Winter mean sea surface temperature (SST) anomalies in the Nino3.4 region of the Pacific Ocean are used to predict the likelihood of summer mean inflow anomalies for Nurek (Figure A1a). When Pacific Ocean SSTs are relatively warm, there is an El Niño phase. This tends to be followed by warmer and wetter winters over Tajikistan. The greater snowpack volume then favours higher than average meltwater flows in the following vegetative season (April to September). Conversely, when Pacific Ocean SSTs are relatively cold, there is a La Niña phase. In this case, winters in Tajikistan tend to colder and drier than average conditions, leading to relatively low snowpack volumes then depressed meltwater flows in the following summer.

Figure A1. Summer inflow anomalies (m^3s^{-1}) to Nurek (at Komsomolabad) linked to the Nino3.4 index in (a) November-December or (b) October-March, further explained by summer air temperature anomalies at Sari-Tash (c,d) for the period 1941-2016. Lines within the box-and-whisker plots show the median; boxes show the interquartile range (IQR); T-bars are 1.5 x IQR; and points are outliers.

(a)   (b)

(c)   (d)

Inflow during the vegetative period also depends on whether the summer in the Vakhsh has relatively cold, average or warm air temperatures. Overall, the lowest mean summer inflows tend to occur after a winter La Niña followed by a cooler than average summer in Tajikistan (Figures A1c and A1d). On the other hand, the highest mean summer inflows tend to occur after a winter El Niño followed by a warmer than average summer. However, the box-and-whisker plots show that there is considerable variation in forecasted flows even within these contrasting phases.

Figures A1a and A1b compare expected summer inflow anomalies based on Pacific Ocean SST anomalies (given by the Nino3.4 index) averaged over November-December and October-March respectively. This shows that differences between the three Nino3.4 phases actually become less marked when the longer (i.e. October-March, Figure A1b) averaging period is used. In other words, there is no benefit in waiting until SSTs at the end of March are known. This means that the best seasonal forecast of inflows for the vegetative season can be provided by late January/early February each year, just after the time when the Nino3.4 signal is often strongest. Table A1 shows the likelihood of different summer inflow anomalies after the three Pacific Ocean phases. Following La Niña or Neutral conditions, it is most likely that summer inflows anomalies will be between 0 to -99 m^3s^{-1} (with a likelihood of 32% for La Niña, and 41% for Neutral conditions). Following El Niño conditions, summer inflow anomalies will most likely fall between +1 to +100 m^3s^{-1} (with 39% likelihood).

Table A1. Likelihood (%) of mean summer (April-September) Nurek inflow anomalies (m^{3}s^{-1}) at Komsomolabad depending on previous **NOVEMBER-DECEMBER** Nino 3.4 phase for the period 1941-2016 (1993-1999 data are missing). Nino3.4 categories are based on 1982-2010 terciles.

Inflow anomaly (m^{3}s^{-1}) at Komsomolabad	Category of inflow anomaly	November-December Nino3.4 index		
		La Niña (Nino3.4 <-0.60°C)	Neutral (Nino3.4 in range -0.60°C to +0.62°C)	El Niño (Nino3.4 >+0.62°C)
below -200	Below normal	10%	0%	0%
-199 to -100		29%	15%	0%
-99 to 0	Near normal	32%	41%	14%
+1 to +100		24%	22%	39%
+101 to +200	Above normal	5%	15%	33%
above +200		0%	7%	14%

Table A1 can also be used to forecast the likelihood of more extreme conditions. For example, following La Niña there is 10% likelihood that summer inflows are more than 200 m^{3}s^{-1} below average (Table A1, top row). Even after El Niño conditions there is 14% likelihood that inflows could be 0 to -99 m^{3}s^{-1} below average (Table A1, third row). Table A2 gives a sense of the uncertainty in inflow forecasts between different phases of the winter Nino3.4 index and within warm to cold summers over the Vakhsh basin. Depending on the winter SST anomaly, expected summer inflow anomalies could range between -276 m^{3}s^{-1} (La Niña, lower bound) and +277 m^{3}s^{-1} (El Niño, upper bound). During a warm summer in the Vakhsh, inflow anomalies could vary between -164 m^{3}s^{-1} to +277 m^{3}s^{-1} (Table A2, top row); in a cold summer the range is -276 m^{3}s^{-1} to +238 m^{3}s^{-1} (Table A2, bottom row).

Table A2. Mean summer (April-September) inflow anomaly (m^3s^{-1}) for Nurek depending on previous **NOVEMBER-DECEMBER** Nino 3.4 SST anomaly in the Pacific and following **APRIL-SEPTEMBER** air temperature anomalies in the upper Vakhsh basin (Sari-Kash station in Kyrgyzstan) during the period 1941-2016. Nino3.4 terciles were based on the period 1982-2010. The data are the same as in Figure A1a and A1c.

Summer air temperature anomaly at Sari Tash (in the upper Vakhsh basin)	November-December Nino3.4 index[31]		
	La Nina (Nino3.4 <-0.60°C)	**Neutral** (Nino3.4 in range -0.60°C to +0.62°C)	**El Nino** (Nino3.4 >+0.62°C)
Above normal	-164 to +105 m^3s^{-1}	-91 to +223 m^3s^{-1}	-4 to +277 m^3s^{-1}
Normal	-165 to +77 m^3s^{-1}	-187 to +204 m^3s^{-1}	-68 to +219 m^3s^{-1}
Below normal	-276 to +8 m^3s^{-1}	-153 to +178 m^3s^{-1}	+6 to +238 m^3s^{-1}

Table A2 captures the uncertainty due to the relatively small number of recorded events in each category, as well as the influence of summer air temperatures in Tajikistan and other unrepresented factors that affect summer inflow forecasts (such as data errors, or local climate conditions).

Note that a baseline period 1982-2010 was used to derive the bounding values (terciles) of the three categories of Nino3.4 index and air temperature in both Tables A1 and A2. This baseline period is the same as the International Research Institute (IRI) seasonal climate forecasts, thereby enabling use of their temperature outlooks to establish whether the summer temperature is likely to be below normal, normal, or above normal (see *Step 6* below).

---

[31] Tabulated values and terciles were based on HadISST Nino3.4. To convert Nino3.4 from OISSTv2 into HadISST apply the correction: HadISST = (OISSTv2)*0.93

10-step application of the model. *The worked example is shown in italics.*

Forecasts can be issued at TWO times, first in January then in April each year:

Seasonal flow forecast issued in late January/early February

Step 1: Access Nino3.4 SST anomalies from the US Climate Prediction Center:

http://www.cpc.ncep.noaa.gov/data/indices/sstoi.indices.

An example of the data format for OISSTv2 is shown in Figure A2 for the period January 1982 to June 1984.

```
YR MON  NINO1+2  ANOM   NINO3   ANOM   NINO4   ANOM NINO3.4   ANOM
1982   1   24.29  -0.17  25.87   0.24   28.30   0.00  26.72    0.15
1982   2   25.49  -0.58  26.38   0.01   28.21   0.11  26.70   -0.02
1982   3   25.21  -1.31  26.98  -0.16   28.41   0.22  27.20   -0.02
1982   4   24.50  -0.97  27.68   0.18   28.92   0.42  28.02    0.24
1982   5   23.97  -0.23  27.79   0.71   29.49   0.70  28.54    0.69
1982   6   22.89   0.07  27.46   1.03   29.76   0.92  28.75    1.10
1982   7   22.47   0.87  26.44   0.82   29.38   0.58  28.10    0.88
1982   8   21.75   1.10  26.15   1.16   29.04   0.36  27.93    1.11
1982   9   21.80   1.44  26.52   1.67   29.16   0.47  28.11    1.39
1982  10   22.94   2.12  27.11   2.19   29.38   0.72  28.64    1.95
1982  11   24.59   3.00  27.62   2.64   29.23   0.60  28.81    2.16
1982  12   26.13   3.34  28.39   3.25   29.15   0.66  29.21    2.64
1983   1   27.42   2.96  28.92   3.29   29.00   0.70  29.36    2.79
1983   2   28.09   2.02  28.92   2.55   28.79   0.69  29.13    2.41
1983   3   28.68   2.16  29.10   1.96   28.76   0.57  29.03    1.81
1983   4   28.56   3.09  29.12   1.62   28.85   0.35  28.91    1.13
1983   5   28.19   3.99  28.97   1.89   29.08   0.29  28.89    1.04
1983   6   27.44   4.62  28.15   1.72   28.88   0.04  28.24    0.59
1983   7   25.95   4.35  26.62   1.00   28.65  -0.15  27.07   -0.15
1983   8   23.78   3.13  25.87   0.88   28.38  -0.30  26.53   -0.29
1983   9   22.24   1.88  25.24   0.39   28.23  -0.46  26.44   -0.28
1983  10   21.86   1.04  24.61  -0.31   27.75  -0.91  25.87   -0.82
1983  11   21.90   0.31  24.17  -0.81   27.76  -0.87  25.58   -1.07
1983  12   23.01   0.22  24.44  -0.70   27.82  -0.67  25.59   -0.98
1984   1   24.18  -0.28  24.82  -0.81   27.64  -0.66  25.64   -0.93
1984   2   25.18  -0.89  26.22  -0.15   27.25  -0.85  26.39   -0.33
1984   3   26.00  -0.52  27.12  -0.02   27.21  -0.98  26.86   -0.36
1984   4   25.16  -0.31  27.34  -0.16   27.70  -0.80  27.39   -0.39
1984   5   23.23  -0.97  26.46  -0.62   27.95  -0.84  27.39   -0.46
1984   6   21.96  -0.86  25.38  -1.05   28.13  -0.71  26.86   -0.79
```

Figure A2. An extract of monthly OISSTv2 Nino indices for the period January 1982 to June 1984. Data for November-December 1983 are highlighted. Source: US Climate Prediction Center.

Step 2: Calculate the mean Nino3.4 anomaly for November ('MON' 11) and December ('MON' 12) using data in the last column of Figure A2 ('ANOM').

For example, in 1983 the mean Nino3.4 anomaly was -1.03 °C (that is -1.07 °C in November plus -0.98 °C in December, divided by 2). In this case, because the mean anomaly was -1.03 °C, there were cooler than average SSTs in the Nino3.4 region of the Pacific Ocean.

Step 3: Compare the November-December mean Nino3.4 index from *Step 2* with the tercile boundary values shown at the top of each column in Table A1.

In the example of 1983, -1.07 °C is less than -0.60 °C, so the winter falls into the La Niña category – the central Pacific Ocean is cooler than average.

Step 4: Based on the Nino3.4 phase established in *Step 3*, refer to the appropriate column and rows of inflow anomalies in Table A1 to derive the likelihood of above, near normal or below normal inflows.

For the case of 1983, with winter La Niña conditions, in summer 1984 the likelihood of: (a) below normal inflows would have been 39% (i.e. 10% + 29% from Table A1); (b) near normal inflows would have been 56% (i.e. 32% + 24% from Table A1); and (c) above normal inflows would have been 5% (i.e. 5% + 0% from Table A1).

Step 5: Issue the January forecast of expected summer inflows to decision-makers. Be clear that a seasonal forecast can never be 100% certain about the expected inflow anomaly, so the outlook should only be used to inform (not dictate) any operational decisions or contingency measures.

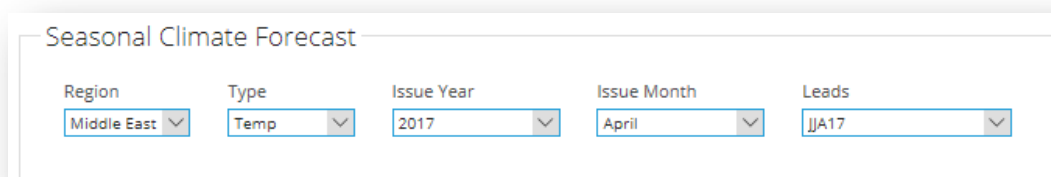*The observed mean inflow in summer 1984 was actually above normal which, as noted above, only had 5% likelihood.*

Seasonal forecast update in April

Step 6: Refer to the International Research Institute (IRI) summer temperature forecast for the Middle East region issued in April:
http://iri.columbia.edu/our-expertise/climate/forecasts/seasonal-climate-forecasts/.
Select the 'Region' as 'Middle East'; 'Type' as 'Temp'; 'Issue Year' as the current year; 'Issue Month' as 'April and 'Leads' as 'JJA' (June/ July/ August) (Figure A3).

Figure A3. Drop down menus for IRI seasonal climate forecast issued in April. Source: IRI.

A map of forecasted temperatures like that shown in Figure A4 is produced.

Note down the most likely summer (June / July / August) temperature conditions for the Vakhsh basin above Nurek, i.e. 'Below Normal', 'Normal' or 'Above Normal'.

For this example, let us assume that the IRI temperature forecast for the Vakhsh basin above Nurek in summer 1984 was "Above Normal" – as shown in Figure A4. (We have made this assumption for illustrative purposes because IRI temperature forecasts are not available as far back as 1984).
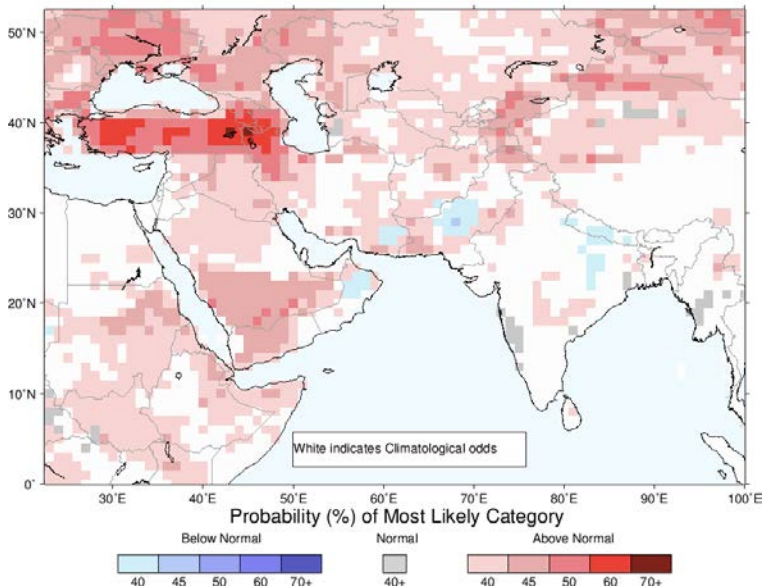

Figure A4. An example IRI seasonal forecast for the Middle East showing expected air temperature anomalies for Tajikistan. The three forecast categories (i.e. Below Normal, Normal, and Above Normal) each have 33% likelihood due to chance alone. The example forecast shows that the most likely category for Tajikistan on this occasion was 'Above Normal'. For this example forecast, in the eastern half of the country there is 50% likelihood of Above Normal temperatures (i.e. greater than the 33% expected by chance).

Step 7: Refer to Table A2 which gives the range of inflow anomalies depending on both the Nino3.4 SST anomaly **and** expected summer air temperature anomaly for Tajikistan. Use the output from *Step 3* (i.e. the mean Nino3.4 anomaly) with *Step 6* (i.e. the IRI temperature forecast for the Vakhsh basin) to select the correct cell in Table A2.

As noted before, the winter of 1983/84 falls into the La Niña category, because the Nino3.4 SST anomaly of -1.03 °C is less than -0.60 °C (see the tercile boundary values at the top of each column in either Table A1 or A2).

Using Table A2, the range of expected mean summer inflow anomalies following La Niña conditions in November-December, with forecasted "Above Normal" summer temperatures for the Vakhsh basin lies between -164 m^3s^{-1} and +105 m^3s^{-1} (Table

Step 8: Compare the expected range of inflows derived from *Step 7* with the summer mean inflow forecast produced by Tajik Hydromet (which is also issued in early April). Their forecast is based on a regression relationship between winter half-year (October-March) total precipitation in the upper Vakhsh basin and the following summer half-year (April-September) discharge entering Nurek.

Step 9: Issue a consensus outlook of summer mean inflows to Nurek based on all available information (i.e. the output of *Step 7* alongside the Tajik Hydromet seasonal inflow forecast, *Step 8*, and any snowpack survey data).

Step 10: After the summer, compare both forecasts with the observed flow. Update the archive of forecasted and observed flows then publish skill metrics (such as the Heidke Hit Proportion, see #1 below) for both forecast techniques.

Advanced technical notes

1. The statistical model presented in this Technical Annex has a forecast success rate of 51-59% depending on the period of record. This measure (the Heidke Hit Proportion) describes the frequency with which the correct inflow anomaly category is predicted (i.e. below average, average, or above average). A forecast skill of 33% would be expected by chance alone. Decision-makers must judge whether this level of significant statistical skill (i.e. 51-59%) has practical significance.

2. The summer mean inflow anomalies in Tables A1 and A2 are sensitive to five factors: (1) the period of discharge data available to compute anomalies (here 1941-2016, with 1993-1999 missing); (2) the baseline period used to calculate the inflow anomalies (here 1982-2010); (3) the baseline period used to calculate the Nino3.4 tercile boundaries (here 1982-2010); (4) the baseline period used to calculate summer mean air temperature tercile boundaries (here 1982-2010); (5) the temperature record used to represent the Vakhsh basin (here Sari-Tash).

3. Observed inflow data were available for two stations on the River Vakhsh: Garm (1941-1986, 1988 and 1990) and Komsomolabad (1949-1957, 1977-1992, 2000-2016). A blended summer mean inflow series was created for Komsomolabad by regressing against flows at Garm using overlapping periods of record (1949-1957 and 1977-1990) (Figure A4). The model was tested for biases due to non-stationarity and autocorrelation. Inflow anomalies were derived with respect to the mean summer inflow of the blended series during the period 1982-2010.
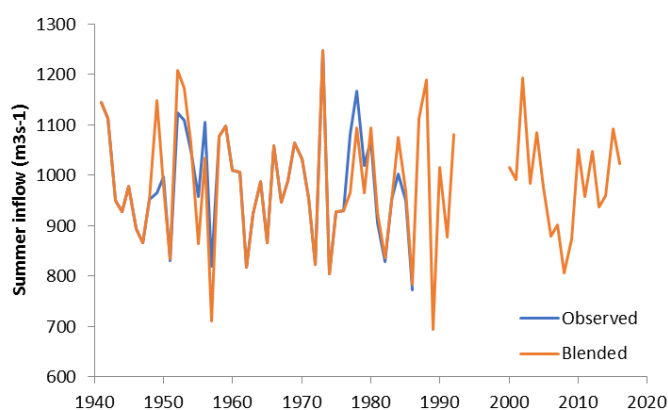

Figure A4. Observed and blended inflows for Komsomolabad 1941-2016 (1993-1999 missing). The regression model explains 68% of observed variation in summer mean flow at Komsomolabad.

4. The Nino3.4 data used to construct Tables A1 to A2 came from the HadISST archive which extends back to the 1880s. Unfortunately, the update time for the publicly available data is too long to be useful for seasonal flow forecasting in Tajikistan. Hence, the OISSTv2 Nino3.4 index of the Climate Prediction Center is

recommended. This is updated every month. Although HadISST and OISSTv2 are very similar, for absolute accuracy the latter can be converted to HadISST Nino3.4 using a correction factor 0.93 (see footnote #1). This only makes a difference if the OISSTv2 value is close to a Nino3.4 phase boundary (Table A1).

5. Mean summer air temperatures for Sari-Tash were used, rather than Lyakhsh. The former station is at higher elevation and, therefore, better represents the snowpack melt conditions in the headwaters of the Vakhsh basin.

6. The uncertainty range for summer inflow anomalies in Tables A1 and A2 partly reflects carryover or depletion of snowpack between successive winters. For instance, larger than average melt in one summer could mean that the initial snowpack store in the following winter starts from a low volume. Even if average amounts of winter precipitation occurs, the likelihood of higher than average summer flows in the next year is reduced. A late March snowpack survey could be a more reliable guide than temperature for determining whether the expected summer flow will be at the high or low end of the ranges in Table A2.

7. Use of recent discharge data alongside records from the 1940s and 1950s includes any long-term climate change such as increased contributions to inflows from glacier melt and/or extended melt seasons outside the assumed vegetative season (e.g. now potentially spanning March to October). These factors could introduce non-stationarities into the model such that summer inflow anomalies for the present are underestimated when earlier data reflecting cooler ambient conditions are included. This possibility warrants further research into the influence of discharge data period on the expected range of inflow anomalies.

8. In view of point #3 – and the incremental lengthening of the available discharge record with each passing year – values in the look-up tables should be periodically reviewed and refreshed. Ideally, this would not happen annually, to avoid confusion, but every 5-10 years depending on the pace of hydroclimatic change and occurrence of any notable extreme events in the region.

9. The statistical model described in this Technical Annex could be further developed and modified to forecast: (a) winter air temperatures/ heating degree days (to predict winter energy demand); (b) spring geohazard risk from avalanches, rockfalls, landslides, flash floods and mudflows (to predict likelihood of damage to infrastructure); (c) summer 10 day maximum inflows (to predict reservoir spilling volumes). This assumes data are available to calibrate the models and that statistically significant, as well as practically significant links can be established with dominant climate modes (e.g. El Niño Southern Oscillation [ENSO], North Atlantic Oscillation [NAO], Pacific Decadal Oscillation [PDO] or Indian Ocean Dipole [IOD]).

10. More generally, a deeper understanding of regional variations in hydroclimatic responses to climate modes could help to identify possible hedging of transboundary water and electricity exports/imports. For instance, electricity trading agreements might be structured such that Tajikistan can export more energy

following El Niño phases (when reservoir water levels are expected to be higher) to neighbours that might be experiencing lower energy production. Conversely, during La Niña, more electricity might be generated by neighbouring countries and imported to Tajikistan to bolster winter supplies.

Reference

Dixon, S.G. and Wilby, R.L. 2017. Potential for seasonal forecasting of reservoir inflows in Central Asia. *Environmental Research Letters*, under revision.