HYDROMETEOROLOGICAL VARIABILITY OVER PAKISTAN

by

Furrukh Bashir

Copyright © Furrukh Bashir 2017

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF HYDROLOGY AND ATMOSPHERIC SCIENCES

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

WITH A MAJOR IN HYDROMETEOROLOGY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2017

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation prepared by *Furrukh Bashir*, titled *HYDROMETEOROLOGICAL VARIABILITY OVER PAKISTAN*

and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy

Date: (November 17, 2017)
Hoshin V. Gupta

Date: (November 17, 2017)
Xubin Zeng

Date: (November 17, 2017)
Thomas J. Galarneau

Date: (November 17, 2017)
Pieter Hazenberg

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

Date: (November 17, 2017)
Dissertation Co-Director: Hoshin V. Gupta

Date: (November 17, 2017)
Dissertation Co-Director: Xubin Zeng

The University of Arizona

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of the requirements for an advanced degree at the University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that an accurate acknowledgement of the source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: Furrukh Bashir

ACKNOWLEDGEMENTS

I gratefully acknowledge efforts made by Senator James William Fulbright that culminated as establishment of an international exchange program that allowed me to pursue education and research at Department of Hydrology and Atmospheric Sciences, University of Arizona, that would help me to perform as a better scientist in field of hydrometeorology.

I acknowledge the efforts made by United States Education Foundation Pakistan (USEFP), and its head Rita Akhtar, who are trying to educate Pakistani youth and building the capacity of the professional, tirelessly for past several years.

I am grateful to Hoshin V. Gupta, Xubin Zeng, James Shuttleworth, Francina Dominguez, Aveline F. Arellano, Paul Sheppard, Elizabeth Ritchie, Christopher Castro, Chuck Weidman, Pieter Hazenberg, Thomas J. Galarneau and Eric A. Betterton, who spent a lot of their personal time with me and kept me motivated throughout this period and never gave up on me.

I would say thank you to my fellow students and to administrative and supporting staff of Department of Hydrology and Atmospheric Sciences, especially, to Sandy Holford, Lupe Romero, Sarah Warren, Mike Eklund and Matthew Jones for caring attitude and continuous support.

I would say special thanks to Sarah McCormick, Stephanie Sasz, Betsy Barringer and Brian Diffley at Institute of International Education (IIE), and to Joanne Lagasse- Long at International Student Services who were always available, to me in time of need.

I am thankful to colleagues at Pakistan Meteorological Department who kept on supporting me, throughout my career.

Finally, I am grateful to my mother and wife who endured my prolonged absence from the country and managed chores on my proxy, and I would say thanks to Abeera and Adil for being part of my life.

TABLE OF CONTENTS

LIST OF TABLES 8
LIST OF FIGURES 9
ABSTRACT 13
CHAPTER 1: INTRODUCTION 15
1.1 THE KARAKORAM ANOMALY 15
1.2 AVAILABILITY OF HYDROMETEOROLOGICAL DATA 17
1.3 ASSESSMENT OF DROUGHTS IN PAKISTAN 17
1.4 OBJECTIVES 19
CHAPTER 2: PRESENT STUDY 20
2.1 A HYDROMETEOROLOGICAL PERSPECTIVE ON THE KARAKORAM ANOMALY USING UNIQUE VALLEY-BASED SYNOPTIC WEATHER OBSERVATIONS 20
2.2 DEVELOPMENT OF HIGH RESOLUTION DAILY PRECIPITATION AND TEMPERATURE DATA PRODUCT FOR PAKISTAN 20
2.3 RECONSTRUCTION OF HISTORY OF METEOROLOGICAL AND HYDROLOGICAL DROUGHTS IN PAKISTAN USING HIGH RESOLUTION IN SITU GRIDDED DATA SETS 21
2.4 FUTURE RESEARCH DIRECTIONS 21
REFERENCES 23
APPENDIX A: A HYDROMETEOROLOGICAL PERSPECTIVE ON THE KARAKORAM ANOMALY USING UNIQUE VALLEY-BASED SYNOPTIC WEATHER OBSERVATIONS 29
ABSTRACT 30
1. INTRODUCTION 30
2. DATA AND METHODS 32
3. RESULTS 34
3.1 SEASONAL AND ANNUAL HYDROLOGIC BUDGET 34
3.2 HUMIDITY AND CLOUD COVER 35
3.3 WIND SPEED 36
3.4 NET RADIATION 37

3.5 POTENTIAL EVAPOTRANSPIRATION AND CLIMATE MOISTURE INDEX (CMI) 38
4. DISCUSSION 38
5. CONCLUSION 40
ACKNOWLEDGMENTS, SAMPLES, AND DATA 40
REFERENCES 41
SUPPORTING INFORMATION 52
APPENDIX B: DEVELOPMENT OF HIGH RESOLUTION DAILY PRECIPITATION AND TEMPERATURE DATA PRODUCT FOR PAKISTAN 78
ABSTRACT 79
1. INTRODUCTION 79
2. STUDY AREA AND CLIMATOLOGY 82
2.1 STUDY AREA 82
2.2 CLIMATOLOGY OF PAKISTAN 83
3. GENERATION OF A GRIDDED DATA PRODUCT FOR PAKISTAN 84
3.1 PRECIPITATION AND TEMPERATURE DATA 84
3.2 ACCOUNTING FOR MISSING DATA 85
3.3 CROSS-VALIDATION OF ICDWM IMPUTATION 86
3.4 DEVELOPMENT OF THE GRIDDED CLIMATOLOGY AND DAILY GRIDDED DATA PRODUCT 87
3.5 SENSITIVITY OF THE INTERPOLATION ALGORITHM TO OBSERVATION DENSITY 88
4. RESULTS AND DISCUSSION 89
4.1 MISSING DATA IMPUTATION 89
4.2 COMPARISON OF PAK-HYM-1.0 WITH OTHER PRODUCTS 92
4.3 PERFORMANCE OF THE INTERPOLATION ALGORITHM 94
5. CONCLUSIONS 96
ACKNOWLEDGMENTS, SAMPLES, AND DATA 98
REFERENCES 98

SUPPORTING INFORMATION 113

APPENDIX C: AN INVESTIGATION AND ASSESSMENT OF METEOROLOGICAL AND HYDROLOGICAL DROUGHTS IN PAKISTAN USING HIGH RESOLUTION IN SITU GRIDDED DATA SET 126

ABSTRACT 127

1. INTRODUCTION 127
2. DATA AND STUDY AREA 130
   - 2.1 HIGH RESOLUTION GRIDDED HYDROMETEOROLOGY DATA FOR PAKISTAN 131
   - 2.2 RIVER INFLOWS DATA 131
3. METHODS 132
   - 3.1 STANDARDIZED PRECIPITATION INDEX (SPI) 132
   - 3.2 KOLMOGOROV-SMIRNOV GOODNESS OF FIT TEST 133
   - 3.3 MAXIMUM COVARIANCE ANALYSIS 134
     - 3.3.1 PERIODOGRAM 135
   - 3.4 STANDARDIZED RUNOFF INDEX (SRI) 135
   - 3.5 PALMER DROUGHT SENSITIVITY INDEX (PDSI) 136
4. RESULTS AND DISCUSSION 138
5. CONCLUSIONS 142

ACKNOWLEDGMENTS, SAMPLES, AND DATA 144

REFERENCES 145

SUPPORTING INFORMATION 155

LIST OF TABLES

Table A. 1 a) List of meteorological sites used in the study with their latitude, longitude, elevation (meters) and description of the data used in the study. Symbol of ☀ indicates that monthly mean of maximum, mean and minimum temperature from that site is utilized and symbol of ☁ indicates that additional data such as monthly means of synoptic weather data reported by meteorological stations in morning (08:00 a.m., local time) and afternoon (05:00 p.m., local time) along with accumulated precipitation are utilized for the study. b) List of observed and derived meteorological variables utilized to perform analyses. 48

Table B. 1: Estimates of desirable numbers of gauges for different parts of Pakistan, based on geo-statistical analysis of climatological summer precipitation (See Figure B. 7). 105

Table C. 1. Names, gauging location and duration of data of different rivers used in the study. 147

LIST OF FIGURES

Figure A. 1. Study Area and Hydrological Budget. a) Map showing the location of HKH centered in Northern Pakistan, and the meteorological sites used in this study. The symbols of (🌞) and (🌅) are explained in Table 1 and (∗) indicates Tarbela Dam's location. Elevation from mean sea level is presented in legend. Right panels show the accumulated percentage (shaded, left axis) of seasonal contributions in mean regional annual precipitation (1961-2011), b), and discharge recorded at Tarbela Dam (April 1961-March 2005), c), while red line in (b) indicates regional mean of annual accumulated precipitation (mm, right axis) and in (c) it indicates the total annual sum of specific runoff (mm) at Tarbela Dam. The legend brackets indicate trends in seasonal contributions (from winter to autumn, in % per decade) to the mean regional precipitation (b) and discharge (c), respectively. Statistical significance (95% confidence interval) of the trends is marked with an asterisk sign. 49

Figure A. 2. Bar plots (on both sides) indicate mean regional trends of the seasonal means of different meteorological variables with one standard deviation on both sides of the bar indicated by the horizontal line; from top to bottom are winter, spring, summer, autumn and annual, respectively. Time series plots show the corresponding summer anomalies standardized with respect to mean and standard deviation for 1961-1990. This base period is selected to present change in variables with respect to Pre-Karakoram Anomaly conditions. a) maximum temperature; b) mean temperature; c) minimum temperature; d) diurnal temperature range; e) dry bulb temperature at 08:00 a.m.; f) dry bulb temperature at 05:00 p.m.; g) vapor pressure at 08:00 a.m.; h) vapor pressure at 05:00 p.m.; i) cloud cover at 08:00 a.m.; j) cloud cover at 05:00 p.m.; k) near surface wind speed at 08:00 a.m.; l) near surface wind speed at 05:00 p.m.; m) surface pressure at 08:00 a.m.; n) surface pressure at 05:00 p.m.; o) potential evapotranspiration; p) accumulated precipitation; q) net radiation; r) climate moisture index. 50

Figure B. 1. Hypsometry and climatology of Pakistan: Map presenting elevation of Pakistan (Center) along with 365-day climatologies of maximum temperature (red line), minimum temperature (blue line) and precipitation (black line) at various locations of the country (subplots 1-12). Each bold line indicates the 365-day smoothed climatology for a specific location and the filled region indicates the 25th to 75th percentile spread. The y-axis indicates temperature

($°\boldsymbol{C}$) on odd numbered subplots and precipitation ($\boldsymbol{mm/day}$) on even numbered subplots.....106

Figure B. 2. Taylor Diagrams presenting spatial cross-validation of ICDWM imputation in time series for year 2011 via the leave-one-out method; (a) Mean Precipitation, (b) Mean Maximum Temperature and (c) Mean Minimum Temperature, all at monthly scales. Each plot presents the centered root mean square error (CRMSE) on the x-axis, normalized standard deviation on the y-axis, and correlation between observations and reconstructed time series along the circumference of the quarter circle. See section 3.3 for details.....107

Figure B. 3. Bar and whisker plots indicating the probability of type I, type II, and type III events for (a) low intensity CAT-I, (b) moderate intensity CAT-II, and (c) extreme intensity CAT-III events. Subplots (a1-a3, b1-b3, c1-c3) indicate, at observations sites, the probabilities of event types in each category. See section 4.1 for details. 108

Figure B. 4. Comparative evaluation of PAK-HYM-1.0 mean summer precipitation ($\boldsymbol{mm/day}$) with alternative precipitation products; (a) APHRODITE (1960-2007), (b) TRMM-3b43 (1998-2013), (c) CHIRPS (1981-2013), (d) CMORPH (2002-2013), (e) CRU (1961-2013), (f) CPC (1979-2013), (g) GPCC (1960-2013), (h) PERSIANN (2000-2013), (i) PAK-HYM-1.0 (1960-2013). Subplot (j) presents a spatial statistical comparison of PAK-HYM-1.0 with other products for the summer season. See section 4.2 for details.....109

Figure B. 5. Comparison of PAK-HYM-1.0 with APHRODITE, CHIRPS and CMORPH precipitation products for individual events at the monthly and daily scales; (a & b) Monthly comparison with APHRODITE for July 1992, (c & d) Daily comparison with APHRODITE for July 24, 1992, (e & f) Monthly comparison with CHIRPS for June 2010, (g & h) Daily comparison with CMORPH for Aug 03, 2010. See section 4.2 for details. 110

Figure B. 6. Sensitivity of the interpolation algorithm to number of observations used. The y-axis shows RMSE for multiple realizations and x-axis shows observational density; (a) Mean annual precipitation, (b) Mean annual maximum temperature, (c) Mean annual minimum temperature. See section 4.3 for details.....111

Figure B. 7. The first two columns show experimental semi-variograms for mean summer and winter precipitation in Pakistan with directional anisotropy; (a1)

Summer N-S, (a2) Summer E-W, (b1) Winter N-S, (b2) Winter E-W. The dashed and solid black lines indicate the semi-variance ($\boldsymbol{\gamma}$) and covariance ($\boldsymbol{C}$), respectively. The dotted red, blue and black lines indicate $\mathbf{0}$, the nugget and the sill, respectively. The last three columns show observational site decorrelation lengths for different numbers of sites; (a3) summer above 35°N, (b3) winter above 35°N, (a4) summer between 30°N and 35°N, (b4) winter between 30°N and 35°N, (a5) summer below 30°N, (b5) winter below 30°N. The bar and whisker plots indicate uncertainties associated with the de-correlation lengths (y-axis) estimated from random subsets of the observations. See section 3.5 and 4.3 for details. 112

Figure C. 1. Hypsometry of Pakistan with the location of meteorological observatories (a), annual mean daily precipitation (b), annual mean daily maximum temperature (c) and minimum temperature (d) of Pakistan from PAK-HYM-1.0. 148

Figure C. 2. Map presenting Indus Basin (Blue boundaries) with some of the river inflows gauging locations (Red points) and long-term WMO affiliated weather stations (Blue points). Light color polygon indicates the political boundaries of Pakistan.....149

Figure C. 3. Example of equiprobability transformation from gamma distribution fit on precipitation to the standard normal distribution in SPI framework. Cumulative probability distribution of October-December (OND) Islamabad precipitation observations (blue line), fitted gamma distribution (black like) (a), similar to (a) with cumulative standard normal distribution (black line) in SPI framework (b), Rejection rates of the Kolmogorov–Smirnov goodness-of-fit test for gridded PAK-HYM-1.0 data for 1-month, 3-months, 6-months, 12-months and 24-months at significance levels of 0.01 (red bars) and 0.05 (blue bars).....150

Figure C. 4. Analyses of Standardized Precipitation Index for Pakistan. PC-1 of the maximum covariance analysis for 3-months SPI is presented in (a1), where x-axis presents years and y-axis presents the magnitude of the standardized anomaly, blue color indicates positive anomalies and red color indicates negative anomalies, correlation map of PC-1 and 3-month SPI gridded time series, where color bar presents the strength of correlation (b1), Eigenvalue spectrum of covariance analysis of 3-month SPI with significance indicated by error bar (c1), periodogram of PC-1 of 3-month SPI covariance analysis

indicating spectrum power density at y-axis and frequency at x-axis (d1). Similarly, 6-, 12- and 24- months analysis is presented in successive rows. 151

Figure C. 5. Standardized Runoff Index (SRI) for different rivers of Pakistan. 3-month SRI for River Astore (a1), 6-month SRI (a2), 12-month SRI (c1), 24-month SRI (d1) and spectral analysis indicating power density and frequency (e1). Same analyses for River Hunza at Dainyour (a2, b2, c2, d2, e2), for River Kabul at Nowshehra (a3, b3, c3, d3, e3), for River Indus at Tarbela (a4, b4, c4, d4, e4), for River Jhelum at Mangla (a5, b5, c5, d5, e5), for River Chenab at Marala (a6, b6, c6, d6, e6), for River Ravi at Balloki (a7, b7, c7, d7, e7), and for River Sutlaj at Qila Ganda Singhwala (a8, b8, c8, d8, e8). Panel e1, e2, e3 and e4 has similar y-axis presented in e1, whereas, e5, e6, e7 and e8 have similar y-axis presented in e5. 152

Figure C. 6. Palmer Drought Sensitivity Index (PDSI) for different location of Pakistan for the period of 1960-2013 is presented. 153

Figure C. 7. Palmer Drought Sensitivity Index map of Pakistan for July 2001. 154

ABSTRACT

Pakistan, as an agriculture based economy, is vulnerable to various hydrometeorological hazards ranging from tropical cyclones, thunderstorms, tornadoes, drought, rain, hail, snow, lightning, fog, wind, temperature extremes, air pollution, and climatic change. However, three of the most pressing challenges in terms of water resource availability, that are different in nature, but are inter-linked to each other are discussed over here.

We begin with the Karakoram Anomaly that is considered as one of the most mysterious and most speculated phenomena on Planet Earth. Though, it is confined to the glaciers in the eastern Hindukush, western Karakoram and northwestern Himalayan mountain ranges of Northern Pakistan that are not responding to global warming in the same manner as their counterparts elsewhere, because, their retreat rates are less than the global average, and some are either stable or growing. However, the Karakoram Anomaly has baffled scientific society for more than a decade since its earliest discovery in the year 2005. The reasons of the Karakoram anomaly were mainly associated to physiography of the area and role of climate was considered marginal till now, as climate is influencing glaciers differently all over the globe. Here, for the first time, we present a hydro-meteorological perspective based on five decades of synoptic weather observations collected by the meteorological network of Pakistan. Analysis of this unique data set indicates that increased regional scale humidity, cloud cover, and precipitation, along with decreased net radiation, near-surface wind speed, potential evapotranspiration and river flow, especially during the summer season, represent a substantial change in the energy, mass and momentum fluxes that are facilitating the establishment of the Karakoram Anomaly. In turn, it is influencing the availability of glacier melt in River Indus in summer season.

Secondly, we developed a hydrometeorological data sets for Pakistan as they are extremely important for water related impact studies and future climate change scenarios. Presently, major sources of gridded temperature and precipitation data generation are in-situ observations, satellite retrieved information and outputs from numerical models. However, each

has its own merits and demerits. Among them gridded observed data sets are considered superior if the gauge density is better. Unfortunately, precipitation gauge network of Pakistan is poorly presented in prior gridded products. Therefore, a daily in-situ observation based, 0.05º×0.05º gridded temperature and precipitation data set for Pakistan, for the period of 1960-2013 is developed. It is named as PAK-HYM-1.0, that is an abbreviation of Pakistan and Hydrometeorology, and 1.0 indicates that it is the first version. This data set is developed by utilizing data from 67 meteorological stations of Pakistan. This number of observation sites is 2 to 4 times higher than that used in prior similar products, and this product can be adopted as an operational information product that can be updated on daily basis.

Finally, we focused on meteorological and hydrological droughts in Pakistan. We have reconstructed history of drought in Pakistan using in situ observations based high resolution gridded data through Standardized Precipitation Index (SPI) methodology on different time scales. Furthermore, we have explained the transition of meteorological drought to hydrological drought using river inflows data of large rivers of Pakistan, and explained the sensitivity of different rivers to rainfall and temperature of different seasons. On the basis of this analysis, we have proposed a solution of construction of water reservoirs to tap water resources from northern mountains as inflows from these mountains has potential to perform as a buffer against droughts in low-lying areas of Pakistan. In addition to that, we have demonstrated the potential of Palmer Drought Sensitivity Index (PDSI) as an operational tool for drought monitoring in Pakistan.

CHAPTER 1: INTRODUCTION

Pakistan is an agriculture based economy, that is largely dependent on in-time availability of water resources. Indus River along with its several tributaries is the main source of water availability in Pakistan, as it supports one of the most extensive irrigation network available throughout the world, that enables it to produce surplus food that beefs up food security outside its geographic extent (Amarasinghe *et al.*, 2004). Increase in withdrawal of water for domestic, industrial and irrigations purpose have altered the hydrology of Pakistan. Low inflows in rivers can seriously influence the water quality, as well. With increase in population of the country, it is expected that pressure on water resources will increase due to both expanding urbanization and extensive irrigation (Immerzeel *et al.*, 2010) for agriculture to feed the growing number of people, that can cease the flow of river water (Molle *et al.*, 2010) to reach Arabian Sea, while making it a closed basin.

Pakistan is exposed to all kind of natural hazards ranging from tropical cyclones, thunderstorms, tornadoes, drought, rain, hail, snow, lightning, fog, wind, temperature extremes, air pollution, and climatic change. However, in terms of water resource availability three important challenges, that are different in nature, but are inter-linked to each other are discussed here.

1.1 THE KARAKORAM ANOMALY

Indus basin spans in four countries namely Pakistan, India, China and Afghanistan with its upper portion resting in mountain ranges of Hindukush, Karakoram and Himalaya (HKH). The largest portion of the basin is in Pakistan (52%) followed by India (33%). Almost, all the major rivers in Indus basin are transboundary rivers and are regulated by water rights treaties that enforce abstractions with little regard to environment and ecology. A major portion of water inflows in Indus River and its tributaries is from snow and glacier melt from HKH in spring and summer season. Some studies report up to 61% of the annual discharge in Indus River (above Tarbela Dam) is from snow and glacier melt (Bookhagen and Burbank, 2010). In Pakistan, the extensive irrigation network is regulated mainly by two storage dams namely Tarbela Dam on the

Indus River and the Mangla Dam on the Jhelum River. Indus River, Jhelum River and Kabul River are fed by the upper reaches of the Indus Basin and any change in upstream water supply in these rivers will have a profound effect on several hundred million people living downstream. While most of the glaciers are melting and retreating across the globe, and thereby vulnerable to decrease in their meltwater availability, the glaciers in HKH ranges centered in northern Pakistan are behaving differently by not retreating with global retreat rate or being stable or growing in some cases (Bolch *et al.*, 2012; Fowler and Archer, 2006; Gardelle *et al.*, 2012; Hewitt, 2005, 2007). Hewitt (2005) coined the termed 'The Karakoram Anomaly' that is referred to the expansion of HKH glaciers since 1990s. Previously these glaciers were diminishing since 1920's expect some short term advances in 1970s. Paterson (2000: 358) perceived that Karakoram glaciers are surging as they are residing in relatively young mountains ranges that are undergoing rapid surges. However, Gardelle *et al* (2012) estimated a positive mass balance of $+0.11 \pm 0.22\ myr^{-1}$ water equivalent using regional change in ice elevation by differencing two digital elevation models (DEMs). They found this anomalous gain is in agreement with the reduction in river runoff noted by Fowler and Archer (2006). Till now, the Karakoram Anomaly is not fully resolved, primarily due to inaccessible terrain and extreme environmental conditions that prevent long-term field programs focusing on glacier mass balance. Furthermore, limited long-term near-glacier climate data is available to fully explore the phenomena. Due to poor understanding of the role of climatic forcing in the establishment of the Karakoram Anomaly it is thought to be an outcome of the physiography (Paul *et al.*, 2015) and thickness of widespread debris covering glaciers (Mayer *et al.*, 2006) that is providing them an insulation coating, where climate's role is questionable (Qiu, 2015). On the other hand, a slight gain in mass in Karakoram glacier may give an excuse to some people to deny the existence of global warming due to human activities. Any future climate projection scenario is not applicable to the Karakoram anomaly and future water availability in Upper Indus Basin (UIB), if, climatic causation of the Karakoram Anomaly is not well understood, else it will be mere speculation that has little value for policy making to ensure future food security.

1.2 AVAILABILITY OF HYDROMETEOROLOGICAL DATA

For effective water resource management, it is imperative to deliver credulous hydrometeorological information to concerned resource managers and scientific community. Such information is valuable for water related impact studies, such as, impact of change in rainfall on crop yield or flood scenarios with respect to change in precipitation. Such hydrometeorological information usually requires spatially and temporally continuous data. In the same manner, to identify relationship between large scale circulation and climate of an area the usage of area average climatic indices is highly desirable (Chang *et al.*, 2004; Lau and Wu, 2001; Mariotti, 2007). Such information is available from spatially continuous gridded data sets that are developed using objective analyses of the irregular point observations (Chen *et al.*, 2002, 2008), rather than point observations themselves. Similarly, another important aspect of water resource management is understanding change in precipitation patterns in future climate projections on regional scales that necessitates a direct comparison between the climate models and gridded observations at same spatial scale, as models generally agree to the representative area average rather than the point processes (Osborn and Hulme, 1998). Moreover, operational applications such as climate disaster monitoring by different National Hydro-Meteorological Services (NHMS) also use gridded temperature and precipitation data as an input. Presently, major sources of gridded temperature and precipitation data generation are in-situ observations, satellite retrieved information and outputs from numerical models. However, each has its own merits and demerits (Xie and Arkin, 1997). Among them, gridded observed data sets are considered superior if the gauge density is better. Unfortunately, precipitation gauge network of Pakistan is poorly presented in these gridded products. Therefore, there is a strong need to develop a product using historic in situ hydrometeorological data of Pakistan at high spatial resolution in gridded format that may serve all the aforementioned purposes.

1.3 ASSESSMENT OF DROUGHTS IN PAKISTAN

As already mentioned, agriculture in Pakistan is the main driving force of economic activities as it provides basic ingredients for sustenance and raw material to industry. Increase in

population has created a greater need of growing food that put more pressure on the water resources by increasing its demand. With changing climate and global warming, the agriculture is becoming more vulnerable to climatic hazards, such as, droughts. Though, drought is a recurring feature of climate (Bordi and Sutera, 2007), as it occurs in all climatic zones (Mishra and Singh, 2010), regardless that how much precipitation an area receives climatologically and it is different from aridity that is confined to climatologically low rainfall regions. However, it is related to the reduction in the amount of precipitation received over an extended period, that may range from months to years. Drought affects both surface and groundwater resources and can lead to decrease in water supply, deteriorate water quality, fail agriculture and farming, decrease hydro-power generation, disturb riparian ecology and may halt recreational activities. Moreover, it severely affects transport of nutrients and organic matter associated with movement of water through several pathways. Among all the meteorological hazards, droughts rank the first, and they differ from others in several ways; i) droughts are possibly the most slowly developing hazards, ii) that may last for a very long duration, iii) they are least predictable with accuracy, and iv) their far-reaching impacts may affect the economy beyond the area-experiencing the drought. Unlike floods, hurricanes, earthquakes and tornedoes a drought affects water resources and seldom results in structural damages. For the same reason the quantification of its impact and the provision of relief are far more difficult for drought then for other natural hazards (Wilhite, 2000). Bryant (2005) ranked hazard events considering their characteristics and impacts and found drought as the most hazardous climatic event. Past studies on drought in Pakistan (Adnan *et al.*, 2016; Haroon *et al.*, 2016; Xie *et al.*, 2013) have identified that droughts are wide-spread and often occur simultaneously over large areas. They identified three intensive drought episodes; i) the late 1960s to early 1970s; ii) the middle 1980s, and iii) the late 1990s to early 2000s and they identified a 16-year drought recurrence period. However, none of them has explicitly estimated the intensity and severity of the drought, and did not fully decompose seasonal aspect of the drought in Pakistan, to ascertain severity of drought in different seasons. Further, no past study ever estimated duration and intensity of hydrological droughts in rivers of Pakistan.

Therefore, there is a need to reassess meteorological, hydrological and agricultural droughts in Pakistan.

1.4 OBJECTIVES

We are focusing on three different but inter-related issues that are posing serious challenges to Pakistan in terms of water resources availability and management.

First, we evaluate the role of hydrometeorological forcing in the establishment of the Karakoram Anomaly that is a water resources problem, locally, and one of the most mysterious and insufficiently explained phenomenon, globally.

Further, to satisfy the need of a product that may provide spatially and temporally continuous hydrometeorological information across the country, we develop a daily in-situ observation based, $0.05^{o}\times0.05^{o}$ gridded temperature and precipitation data set for Pakistan that covers the period of 1960-2013 without missing data. It is superior to other similar products as it takes full advantage of meteorological observation network of Pakistan that is poorly presented in other data sets and is upgradeable on daily basis.

Finally, we reconstruct the history of drought in Pakistan using in situ observations based high resolution gridded data through SPI methodology on different time scales. Further, we explain the transition of meteorological drought to hydrological drought and identify the sensitivity of different rivers to the most dominant drought behavior in Pakistan.

CHAPTER 2: PRESENT STUDY

2.1 A HYDROMETEOROLOGICAL PERSPECTIVE ON THE KARAKORAM ANOMALY USING UNIQUE VALLEY-BASED SYNOPTIC WEATHER OBSERVATIONS

First, using hydrometeorological data of synoptic weather observations situated in different valley floors of HKH mountains, along with inflows data of Indus River, we present a detailed analysis that provides an elaborated hydrometeorological perspective of the Karakoram Anomaly. Our finding indicates that energy, mass, and momentum fluxes have been acting to offset each other in increasing and decreasing patterns over the last five decades. Increases in water vapor, cloudiness, and precipitation and decreases in net radiation, near-surface wind speed, and potential evapotranspiration have resulted in a positive hydrologic mass balance at the annual scale. With change in these hydrometeorological variables the HKH watersheds have been turned moisture surplus and energy deficient in recent years. Further, nourishment of the glaciers in both winter and summer has been increasing lately. Overall, our findings clearly support the existence of the Karakoram anomaly due to a positive mass balance of glaciers in the region. It clearly explains the role of climatic factors in the development of Karakoram anomaly by establishing a positive mass balance of glaciers in HKH, along with other non-climatic factors. Our explanation of the Karakoram Anomaly helps to not only address challenges to water resources in Pakistan but also answer a long-standing scientific puzzle that was once considered very hard to answer due to non-availability of very high elevation near glacier hydrometeorological data sets.

2.2 DEVELOPMENT OF HIGH RESOLUTION DAILY PRECIPITATION AND TEMPERATURE DATA PRODUCT FOR PAKISTAN

Hydrometeorological data sets are important for water related impact studies and future climate change scenarios. Presently, major sources of gridded temperature and precipitation data generation are in-situ observations, satellite retrieved information and outputs from numerical models. However, each has its own merits and demerits (Xie and Arkin, 1997). Gridded observed

data sets are considered superior if the gauge density is better. Unfortunately, precipitation gauge network of Pakistan is poorly presented in prior gridded products. Therefore, a daily in-situ observation based, 0.05º×0.05º gridded temperature and precipitation data set for Pakistan, for the period of 1960-2013 is developed here. It is named as PAK-HYM-1.0, that is an abbreviation of Pakistan and Hydrometeorology, and 1.0 indicates that it is the first version. This data set is developed by utilizing data from 67 meteorological observatories of Pakistan. This number of observation sites is 2 to 4 times higher than that used in other similar products.

2.3 RECONSTRUCTION OF HISTORY OF METEOROLOGICAL AND HYDROLOGICAL DROUGHTS IN PAKISTAN USING HIGH RESOLUTION IN SITU GRIDDED DATA SETS

In the third study, we have reconstructed history of drought in Pakistan using in situ observations based high resolution gridded data through SPI methodology on different time scales. Furthermore, we have explained the transition of meteorological drought to hydrological drought using river inflows data of large rivers of Pakistan, and explained the sensitivity of different rivers to rainfall and temperature of different seasons. Based on the analysis, we have proposed a solution of construction of water reservoirs to tap water resources from northern mountains as inflows from these mountains, and this solution has the potential to perform as a buffer against droughts in southern parts of Pakistan. In addition to that, we have demonstrated the potential of PDSI as an operational tool for drought monitoring in Pakistan.

2.4 FUTURE RESEARCH DIRECTIONS

Water resource monitoring and management is multifaceted problem in Pakistan. In perspective of the present study, following areas needs further research:

a) What are the causes of increase in humidity in HKH mountain, especially in summer season. Does it have a direct relation to summer monsoon moisture from

Arabian Sea or is it just recycled evaporated water from plains of intermediate and lower Indus Basin.

b) The decrease in near surface wind speed in HKH is one of the most important factor in establishment of the Karakoram anomaly. It needs further research if this weakening of wind speed is related to northward shift of jet stream or it is related to weakening of local katabatic winds that are driven by the temperature gradient of the glacier surface with ambient air temperature.

c) Increase in surface pressure is observed all over the HKH, especially in summer season. What atmospheric factors can explain it?

d) There is a need to add more hydrometeorological variables in successive version of PAK-HYM. Its spatial coverage can be expanded using data from the neighboring countries to cover whole Indus basin for improving hydrometeorological information on basin scale.

e) Now, the Indus basin is practically closed with near zero environmental flows to Arabian Sea in most of the years (Sharma, 2000), even though it has the potential to receive more than $250\ km^3$ of water each year. Therefore, it is also an interesting research question regarding what fraction of this rainfall is recycled across the Indus basin and what fraction is contributed by the western disturbances and monsoon on average. Moreover, it is imperative to understand what are the environmental consequences for Indus basin for being a closed basin.

f) Though we have presented potential of PDSI as an operational tool for drought monitoring, it needs to be improved using better estimates of the available water content according to the soil type of the area and it has to be calibrated with actual soil moisture observations across the Pakistan for better results.

REFERENCES

Alexandersson, H. (1986), A homogeneity test applied to precipitation data, J. Climatol., 6(6), 661–675, doi:10.1002/joc.3370060607.

Allen, R. G., L. S. Pereira, D. Raes, M. Smith, and W. Ab (1998), Allen_FAO1998, , 1–15.

Archer, C. L., and K. Caldeira (2008), Historical trends in the jet streams, Geophys. Res. Lett., 35(8), 1–6, doi:10.1029/2008GL033614.

Archer, D. (2003), Contrasting hydrological regimes in the upper Indus Basin, J. Hydrol., 274(1–4), 198–210, doi:10.1016/S0022-1694(02)00414-6.

Archer, D. R., and H. J. Fowler (2004), Spatial and temporal variations in precipitation in the Upper Indus Basin, global teleconnections and hydrological implications, Hydrol. Earth Syst. Sci., 8(1), 47–61, doi:10.5194/hess-8-47-2004.

Armstrong, R. L. (2010), The Glaciers of the Hindu Kush-Himalayan Region, ICIMOD (Inernational Cent. Integr. Mt. Dev., 1–16.

Asad, F., H. Zhu, H. Zhang, E. Liang, S. Muhammad, S. Bin Farhan, I. Hussain, M. A. Wazir, M. Ahmed, and J. Esper (2016), Are Karakoram temperatures out of phase compared to hemispheric trends?, Clim. Dyn., 1–10, doi:10.1007/s00382-016-3273-6.

Bhutiyani, M. R., V. S. Kale, and N. J. Pawar (2007), Long-term trends in maximum, minimum and mean annual air temperatures across the Northwestern Himalaya during the twentieth century, Clim. Change, 85(1–2), 159–177, doi:10.1007/s10584-006-9196-1.

Bocchiola, D., G. Diolaiuti, A. Soncini, C. Mihalcea, C. D'Agata, C. Mayer, A. Lambrecht, R. Rosso, and C. Smiraglia (2011), Prediction of future hydrological regimes in poorly gauged high altitude basins: The case study of the upper Indus, Pakistan, Hydrol. Earth Syst. Sci., 15(7), 2059–2075, doi:10.5194/hess-15-2059-2011.

Bolch, T. et al. (2012), The State and Fate of Himalayan Glaciers, Science (80-.)., 336(6079), 310–314, doi:10.1126/science.1215828.

Bollasina, M. A., Y. Ming, and V. Ramaswamy (2011), Anthropogenic Aerosols and the Weakening of the South Asian Summer Monsoon, Science (80-.)., 334(6055), 502–505, doi:10.1126/science.1204994.

van den Broeke, M. R. M. (1997), Structure and diurnal variation of the atmospheric boundary layer over a mid-latitude glacier in summer, Boundary-Layer Meteorol., 83, 183–205, doi:10.1175/1520-0450(1997)036<0763:MHAMBO>2.0.CO;2.

Busing, R. T., L. A. Stephens, and E. E. C. Clebsch (1950), Climate data by elevation in the Great Smoky Mountains: a database and graphical displays for 1947 – 1950 with comparison to long-term data,

Cogley, J. G. (2012), No ice lost in the Karakoram, Nat. Geosci., 5(May), 1–2, doi:10.1038/ngeo1456.

Cogley, J. G. (2016), Glacier shrinkage across High Mountain Asia, Ann. Glaciol., 57(71), 41–49, doi:10.3189/2016AoG71A040.

Cook, E. R., J. G. Palmer, M. Ahmed, C. A. Woodhouse, P. Fenwick, M. U. Zafar, M. Wahab, and N. Khan (2013), Five centuries of Upper Indus River flow from tree rings, J. Hydrol., 486(August 2012), 365–375, doi:10.1016/j.jhydrol.2013.02.004.

Copland, L., T. Sylvestre, M. P. Bishop, J. F. Shroder, Y. B. Seong, L. a. Owen, A. Bush, and U. Kamp (2011), Expanded and Recently Increased Glacier Surging in the Karakoram, Arctic, Antarct. Alp. Res., 43(4), 503–516, doi:10.1657/1938-4246-43.4.503.

Forsythe, N., H. J. Fowler, X.-F. Li, S. Blenkinsop, and D. Pritchard (2017), Karakoram temperature and glacial melt driven by regional atmospheric circulation variability, Nat. Clim. Chang., 7(August), doi:10.1038/nclimate3361.

Fowler, H. J., and D. R. Archer (2006), Conflicting signals of climatic change in the upper Indus Basin, J. Clim., 19(17), 4276–4293, doi:10.1175/JCLI3860.1.

Gardelle, J., E. Berthier, and Y. Arnaud (2012), Slight mass gain of Karakoram glaciers in the early twenty-first century, Nat. Geosci., 5(5), 322–325, doi:10.1038/ngeo1450.

Hasson, S., J. Böhner, and V. Lucarini (2015), Prevailing climatic trends and runoff response from Hindukush–Karakoram–Himalaya, upper Indus basin, Earth Syst. Dyn. Discuss., 6(1), 579–653, doi:10.5194/esdd-6-579-2015.

Hewitt, K. (2005), The Karakoram Anomaly? Glacier Expansion and the “Elevation Effect,” Karakoram Himalaya, Mt. Res. Dev., 25(4), 332–340, doi:10.1659/0276-4741(2005)025[0332:TKAGEA]2.0.CO;2.

Hewitt, K. (2007), Tributary glacier surges: An exceptional concentration at Panmah Glacier, Karakoram Himalaya, J. Glaciol., 53(181), 181–188, doi:10.3189/172756507782202829.

Hewitt, K. (2014), Glaciers of the Karakoram Himalaya Glacial Environments, Processes, Hazards and Resources.

Hewitt, K., C. P. Wake, G. J. Young, and C. David (1989), Hydrological investigations at Biafo Glacier, Karakoram range, Himalaya: an important source of water for the Indus River, Ann. Glaciol., 13, 103–108.

Immerzeel, W. W., L. P. H. van Beek, and M. F. P. Bierkens (2010), Climate change will affect the Asian water towers., Science, 328(5984), 1382–5, doi:10.1126/science.1183188.

Immerzeel, W. W., F. Pellicciotti, and A. B. Shrestha (2012), Glaciers as a Proxy to Quantify the Spatial Distribution of Precipitation in the Hunza Basin, Mt. Res. Dev., 32(1), 30–38, doi:10.1659/MRD-JOURNAL-D-11-00097.1.

Immerzeel, W. W., N. Wanders, A. F. Lutz, J. M. Shea, and M. F. P. Bierkens (2015), Reconciling high-altitude precipitation in the upper Indus basin with glacier mass balances and runoff, Hydrol. Earth Syst. Sci., 19(11), 4673–4687, doi:10.5194/hess-19-4673-2015.

Iqbal, M. A., A. Penas, A. Cano-Ortiz, K. C. Kersebaum, L. Herrero, and S. del R??o (2016), Analysis of recent changes in maximum and minimum temperatures in Pakistan, Atmos. Res., 168, 234–249, doi:10.1016/j.atmosres.2015.09.016.

Kapnick, S. B., T. L. Delworth, M. Ashfaq, S. Malyshev, and P. C. D. Milly (2014), Snowfall less sensitive to warming in Karakoram than in Himalayas due to a unique seasonal cycle, Nat. Geosci., 7(11), 834–840, doi:10.1038/ngeo2269.

Khan, A., K. S. Richards, G. T. Parker, A. McRobie, and B. Mukhopadhyay (2014), How large is the Upper Indus Basin? The pitfalls of auto-delineation using DEMs, J. Hydrol., 509, 442–453, doi:10.1016/j.jhydrol.2013.11.028.

Kozhikkodan Veettil, B., N. Bianchini, U. F. Bremer, É. L. B. Maier, and J. C. Simões (2016), Recent variations of supraglacial lakes on the Baltoro Glacier in the central Karakoram Himalaya and its possible teleconnections with the pacific decadal oscillation, Geocarto Int., 31(1), 109–119, doi:10.1080/10106049.2015.1041565.

Kumar, P. et al. (2015), Response of Karakoram-Himalayan glaciers to climate variability and climatic change: A regional climate model assessment, J. Geophys. Res. Atmos., 118(1), 1–8, doi:10.1002/2015GL063392.Received.

Mayer, C., A. Lambrecht, M. Belò, C. Smiraglia, and G. Diolaiuti (2006), Glaciological characteristics of the ablation zone of Baltoro glacier, Karakoram, Pakistan, Ann. Glaciol., 43, 123–131, doi:10.3189/172756406781812087.

Ming, J., Y. Wang, Z. Du, T. Zhang, W. Guo, C. Xiao, X. Xu, M. Ding, D. Zhang, and W. Yang (2015), Widespread albedo decreasing and induced melting of Himalayan snow and ice in the early 21st century, PLoS One, 10(6), 1–13, doi:10.1371/journal.pone.0126235.

Morgan, K. (2014), Researchers resolve the Karakoram glacier anomaly, a cold case of climate science, Princet. Univ., 39–41.

Mukhopadhyay, B., A. Khan, and R. Gautam (2014), Rising and falling river flows: contrasting signals of climate change and glacier mass balance from the eastern and western Karakoram, Hydrol. Sci. J., 6667(March 2015), 141217125340005, doi:10.1080/02626667.2014.947291.

Nepal, S., and A. B. Shrestha (2015), Impact of climate change on the hydrological regime of the Indus, Ganges and Brahmaputra river basins: a review of the literature, Int. J. Water Resour. Dev., 627(May 2015), 1–18, doi:10.1080/07900627.2015.1030494.

Oerlemans, J., and B. Grisogono (2002), Glacier winds and parameterization of the related surface heat fluxes, Tellus, 54A(5), 440–452, doi:10.1034/j.1600-0870.2002.201398.x.

Qiu, J. (2015), Himalayan ice can fool climate studies, Science (80-.)., 347(6229), 1404–1405, doi:10.1126/science.347.6229.1404.

Quincey, D. J., M. Braun, N. F. Glasser, M. P. Bishop, K. Hewitt, and A. Luckman (2011), Karakoram glacier surge dynamics, Geophys. Res. Lett., 38(18), 2–7, doi:10.1029/2011GL049004.

Quincey, D. J., N. F. Glasser, S. J. Cook, and A. Luckman (2015), Heterogeneity in Karakoram glacier surges, J. Geophys. Res. Earth Surf., 120(7), 1288–1300, doi:10.1002/2015JF003515.Received.

Ragettli, S., W. W. Immerzeel, and F. Pellicciotti (2016), Contrasting climate change impact on river flows from high-altitude catchments in the Himalayan and Andes Mountains., Proc. Natl. Acad. Sci. U. S. A., 113(33), 9222–7, doi:10.1073/pnas.1606526113.

Rankl, M., C. Kienholz, and M. Braun (2014), Glacier changes in the Karakoram region mapped by multimission satellite imagery, Cryosphere, 8(3), 977–989, doi:10.5194/tc-8-977-2014.

Sandeep, S., and R. S. Ajayamohan (2015), Poleward shift in Indian summer monsoon low level jetstream under global warming, Clim. Dyn., 45(1–2), 337–351, doi:10.1007/s00382-014-2261-y.

Seong, Y. B., L. A. Owen, M. P. Bishop, A. Bush, P. Clendon, L. Copland, R. Finkel, U. Kamp, and J. F. Shroder (2007), Quaternary glacial history of the Central Karakoram, Quat. Sci. Rev., 26(25–28), 3384–3405, doi:10.1016/j.quascirev.2007.09.015.

Shuttleworth, W. J. (2012), Terrestrial Hydrometeorology.

Singh, P., and L. Bengtsson (2005), Impact of warmer climate on melt and evaporation for the rainfed, snowfed and glacierfed basins in the Himalayan region, J. Hydrol., 300(1–4), 140–154, doi:10.1016/j.jhydrol.2004.06.005.

Smiraglia, C., C. Mayer, C. Mihalcea, G. Diolaiuti, M. Belò, and G. Vassena (2007), 26 Ongoing variations of Himalayan and Karakoram glaciers as witnesses of global changes: recent studies on selected glaciers, Dev. Earth Surf. Process., 10(6), 235–247, doi:10.1016/S0928-2025(06)10026-7.

Soncini, A., D. Bocchiola, G. Confortola, A. Bianchi, R. Rosso, C. Mayer, A. Lambrecht, E. Palazzi, C. Smiraglia, and G. Diolaiuti (2015), Future Hydrological Regimes in the Upper

Indus Basin: A Case Study from a High-Altitude Glacierized Catchment, J. Hydrometeorol., 16(1), 306–326, doi:10.1175/JHM-D-14-0043.1.

Tahir, A. A., P. Chevallier, Y. Arnaud, and B. Ahmad (2011), Snow cover dynamics and hydrological regime of the Hunza River basin, Karakoram Range, Northern Pakistan, Hydrol. Earth Syst. Sci., 15(7), 2275–2290, doi:10.5194/hess-15-2275-2011.

Tuomenvirta, H. (2002), Homogeneity testing and adjustment of climatic time series in Finland, Geophysica, 38(1–2), 15–41.

Wake, C. P. (1989), Glaciochemical investigations as a tool for determining the spatial and seasonal variation of snow accumulation in the Central Karakoram, northern Pakistan, Ann. Glaciol., 13, 279–284.

WGMS (2012), Fluctuations of Glaciers 2005–2010, World Glacier Monit. Serv. Zurich, Switz., X, 355, doi:10.5904/wgms-fog-2012-11.

Willmott, C., and J. Feddema (1992), A more rational climatic moisture index, Prof. Geogr., 44(1), 84–88, doi:10.1111/j.0033-0124.1992.00084.x.

Winiger, M., M. Gumpert, and H. Yamout (2005), Karakorum-Hindukush-western Himalaya: Assessing high-altitude water resources, Hydrol. Process., 19(12), 2329–2338, doi:10.1002/hyp.5887.

Xu, M., C. P. Chang, C. Fu, Y. Qi, A. Robock, D. Robinson, and H. M. Zhang (2006), Steady decline of east Asian monsoon winds, 1969-2000: Evidence from direct ground measurements of wind speed, J. Geophys. Res. Atmos., 111(24), 1–8, doi:10.1029/2006JD007337.

Zafar, M. U., M. Ahmed, M. P. Rao, B. M. Buckley, N. Khan, M. Wahab, and J. Palmer (2016), Karakorum temperature out of phase with hemispheric trends for the past five centuries, Clim. Dyn., 46(5–6), 1943–1952, doi:10.1007/s00382-015-2685-z.

APPENDIX A: A HYDROMETEOROLOGICAL PERSPECTIVE ON THE KARAKORAM ANOMALY USING UNIQUE VALLEY-BASED SYNOPTIC WEATHER OBSERVATIONS

(Published in Geophysical Research Letters)

Furrukh Bashir
Department of Hydrology and Atmospheric Sciences, University of Arizona, Tucson, AZ, USA
Pakistan Meteorological Department, Islamabad, Pakistan

Xubin Zeng, Hoshin Gupta, Pieter Hazenberg
Department of Hydrology and Atmospheric Sciences, University of Arizona, Tucson, AZ, USA

Bashir, F., Zeng, X., Gupta, H., & Hazenberg, P. (2017). A hydrometeorological perspective on the Karakoram anomaly using unique valley-based synoptic weather observations. Geophysical Research Letters, 44. https://doi.org/10.1002/2017GL075284

© American Geophysical Union. Used with Permission

Key Points:

- The overall hydrologic change in mass for catchments in the Hindukush-Karakoram-Himalaya region, centered in Northern Pakistan, is positive.
- Changes in precipitation, humidity, cloud, river flow, and wind suggest that this region is becoming moisture surplus and energy deficient.
- Changes in energy, mass and momentum fluxes are facilitating establishment of the Karakoram anomaly.

ABSTRACT

Glaciers in the eastern Hindukush, western Karakoram and northwestern Himalayan mountain ranges of Northern Pakistan are not responding to global warming in the same manner as their counterparts elsewhere. Their retreat rates are less than the global average, and some are either stable or growing. Various investigations have questioned the role of climatic factors in regards to this anomalous behavior, widely referred to as 'The Karakoram Anomaly'. Here, for the first time, we present a hydro-meteorological perspective based on five decades of synoptic weather observations collected by the meteorological network of Pakistan. Analysis of this unique data set indicates that increased regional scale humidity, cloud cover, and precipitation, along with decreased net radiation, near-surface wind speed, potential evapotranspiration and river flow, especially during the summer season, represent a substantial change in the energy, mass and momentum fluxes that are facilitating the establishment of the Karakoram Anomaly.

1. INTRODUCTION

The mountain ranges of the Hindukush, Karakoram and Himalaya (HKH) host the world's largest glaciers outside of the Polar Regions, and are known as the 'The Third Pole' of planet Earth (Bocchiola *et al.*, 2011; Soncini *et al.*, 2015). Large portions of the HKH mountain ranges are in Northern Pakistan (Figure 1a). These mountain ranges serve as water towers (Bocchiola *et al.*, 2011; Immerzeel *et al.*, 2010; Ragettli *et al.*, 2016) from where the River Indus

originates, and they supply irrigation and drinking water to satisfy the needs of several million people (Bolch *et al.*, 2012; Immerzeel *et al.*, 2010).

The Karakoram mountain ranges, in northern Pakistan, are influenced by three distinct weather systems: i) westerly storms that account for two-thirds of the high altitude snow accumulation (Bolch *et al.*, 2012; Hewitt, 2005; Hewitt *et al.*, 1989; Wake, 1989), ii) summer monsoon advances that contribute the remaining one-thirds of the high altitude snow accumulation (Archer, 2003; Seong *et al.*, 2007; Wake, 1989; Winiger *et al.*, 2005), and iii) anticyclonic clear weather conditions that affect the behavior of the former two, and which enable solar radiation to reach and melt the glacier (Hewitt, 2005). Accordingly, these glaciers are intermediate between the "summer accumulation type glaciers" of the Greater Himalayan ranges to the east and the "winter accumulation type glaciers" of the European Alpine ranges to the west (Hewitt, 2005), exhibiting characteristics of both the summer and winter accumulation types (Hewitt, 2014).

Global warming, and the corresponding increase in energy flux towards the Earth surface, has resulted in the retreat of mountain glaciers and continuous loss of ice worldwide (WGMS, 2012). In contrast, some studies in the central Karakoram have reported glacial expansion, or at least stability (Bolch *et al.*, 2012; Fowler and Archer, 2006; Gardelle *et al.*, 2012; Hewitt, 2005, 2007), in the highly elevated watersheds, especially since the late 1990's (Copland *et al.*, 2011; Hewitt, 2005; Soncini *et al.*, 2015); this response is widely known as 'The Karakoram Anomaly' (Hewitt, 2005). So far, the role of climatic factors in the development of the Karakoram anomaly is not well understood (Qiu, 2015).

Existence of this anomalous behavior, particularly in the Karakoram ranges (Armstrong, 2010; Bolch *et al.*, 2012; Cogley, 2012, 2016; Fowler and Archer, 2006; Gardelle *et al.*, 2012; Hewitt, 2007; Immerzeel *et al.*, 2012, 2015; Kapnick *et al.*, 2014; Kozhikkodan Veettil *et al.*, 2016; Kumar *et al.*, 2015; Morgan, 2014; Mukhopadhyay *et al.*, 2014; Quincey *et al.*, 2011; Rankl *et al.*, 2014; Smiraglia *et al.*, 2007; Soncini *et al.*, 2015; Tahir *et al.*, 2011), could perhaps be thought of as a simple matter of "cause" (temperature) and "effect" (melting) (Hewitt, 2014), especially since Karakoram temperatures have remained out of phase with hemispheric

temperature trends for the past five centuries (Zafar *et al.*, 2016). However, interpretation of the Karakoram anomaly has been complicated by the lack of long-term field programs focusing on glacier mass balance, and by a scarcity of up-to-date, near-glacier climate data (Gardelle *et al.*, 2012). This is due mainly to the rugged and inaccessible nature of the terrain (Singh and Bengtsson, 2005) and to extreme weather conditions that limit the ability to conduct regular field campaigns. So far, the availability of long-term in-situ meteorological observations is limited mainly to valley-based sites that are (for the most part) situated away from the glaciers (Immerzeel *et al.*, 2012; Winiger *et al.*, 2005). Consequently, previous explanations of the Karakoram anomaly have been based mainly on analyses of long-term temperature, precipitation, and river flow records (Archer, 2003; Archer and Fowler, 2004; Cook *et al.*, 2013; Forsythe *et al.*, 2017; Fowler and Archer, 2006; Hasson *et al.*, 2015; Tahir *et al.*, 2011; Zafar *et al.*, 2016).

The current work presents a detailed analysis that uses long-term meteorological observations from multiple valley floor sites in northern Pakistan. Many of these observed and derived variables such as dry bulb temperature, near-surface wind speed, cloud cover fraction, surface pressure, relative humidity, vapor pressure, net radiation, potential evapotranspiration (ETp) and Climate Moisture Index (CMI) (see Table 1) have not been analyzed before, and enable a novel assessment of the glacial response to regional scale atmospheric changes.

2. DATA AND METHODS

The study area is comprised of the eastern Hindukush, western Karakoram and northwestern Himalayan mountain ranges that are juxtaposed in the center of the province of Gilgit-Baltistan, situated in Northern Pakistan (Figure A. 1a). We use meteorological data from different valley floor sites (500-2500m, Tables 1a) observed over the past five decades (1961-2011). This dataset includes monthly mean maximum temperature and minimum temperature, and monthly accumulated precipitation, along with monthly mean synoptic weather observations (08-AM and 05-PM local time) of dry bulb temperature, wind speed, cloud cover fraction, surface pressure and relative humidity (Figures A. 1s and S1; Tables A.1b). These data, most of which have not previously been reported or analyzed, enable a quantitative assessment of the impacts of net

radiation, wind speed, humidity, cloud cover, atmospheric pressure and potential evapotranspiration on the magnitudes and directions of change in energy, mass and momentum fluxes that are facilitating the establishment of the Karakoram anomaly.

Monthly averages of the observed variables at individual sites are utilized to compute derived variables such as net radiation, vapor pressure, vapor pressure deficit, ETp and CMI. The regionally averaged seasonal means of these hydro-climatological quantities are utilized to facilitate the analyses (see Figure A. 2).

Incoming shortwave and outgoing longwave radiations are computed using an empirical method as presented in equation (S1-S5) of the Supporting Information (S.I) as suggested by (Shuttleworth, 2012), and ETp is estimated using a surface energy balance method (Allen *et al.*, 1998) based on net radiation and near surface fields such as wind speed and vapor pressure deficit (averages of morning and afternoon observations) as explained in equation (S6) of S.I. Data regarding ETp and precipitation is used to estimate values for the CMI (Willmott and Feddema, 1992), that provides an indication of the degree of dryness of the climate at a given location equation (S7). We obtained radiosonde data from the closest station (Kabul, Afghanistan, around 500 km from Central Karakoram) to verify the characteristic changes in energy and momentum of lower regional troposphere. Together with this, discharge recorded at Tarbela dam (April 1961-March 2005), normalized by area of the Upper Indus Basin at Tarbela Dam i.e., 172,159 km2 (Khan *et al.*, 2014) is utilized to assess the seasonal and annual hydrologic budget.

To demonstrate reliabilities of the time series, the ratios of missing values in the time series of various variables are reported in Tables S1 & S2. Homogeneity of all variables is tested at the annual scale using the Standard Normal Homogeneity Test (SNHT) as proposed by (Alexandersson, 1986; Tuomenvirta, 2002) and potential in-homogeneities are reported in Tables S3-S6. Such in-homogeneities, where found, are attributable to strong environmental influences, rather than changes in instrument location or observation practices. Seasonal trends of each variable from all sites, with statistical significance at the 95% confidence interval, are presented in Tables S7-S23.

3. RESULTS

3.1 SEASONAL AND ANNUAL HYDROLOGIC BUDGET

Records of regional mean annual accumulated precipitation for the period 1961-2005 (same period as the discharge data) indicate an increasing trend (red line in Figure A. 1b); note that precipitation gauges installed at the lower elevations record mainly the liquid portion of precipitation. Further, the higher elevations receive about 5-10 times more precipitation than in the valley floor (Hewitt, 2005; Immerzeel *et al.*, 2012; Mayer *et al.*, 2006; Wake, 1989). At the seasonal scale, winter and summer season contributions to regional mean annual total precipitation have been increasing (18% to 24% and 17% to 24%, respectively), while spring and the autumn season contributions have been decreasing (51% to 40% and 14% to 12%, respectively). Meanwhile, records of annual accumulated discharge in the Indus river at Tarbela Dam indicate an overall decreasing trend (red line in Figure A. 1c), but with increases in annual percentage for the winter (5% to 5.6%), spring (11% to 15.5%) and autumn seasons (16% to 17.2%), and a decrease for the summer season (68% to 61.7%). Together, the annual increase in precipitation and decrease in runoff suggests an overall increase in storage of water in the elevated regions of the watershed, mainly in the form of snow and glaciated ice. This explanation is much more likely than the possibility of anthropogenic consumption of summer season river discharge, as no major diversions or storage developments exist upstream of Tarbela Dam that could account for the 6.3% observed decease in summer season contribution. Similar negative trends for summer season discharge above Pertab Bridge are reported by Hasson et al. [2015].

The increase in annual percentage of winter and autumn discharge can be related to the increase in winter season annual percentage precipitation and increases in winter and autumn daytime temperatures (Figures A. 2a, b). However, the 11% decrease in spring season precipitation is accompanied by a 4.5% increase in discharge, indicating that rising temperatures (Asad *et al.*, 2016; Fowler and Archer, 2006; Hasson *et al.*, 2015; Iqbal *et al.*, 2016) have caused the snowfields to become more vulnerable to melt during the spring (Figures A. 2a, b & f). In summer, the 7% increase in precipitation is accompanied by a 6.3% decrease in discharge,

indicating that the regional hydrological mass balance is actually positive during the summer, in spite of the fact that this is when most of the annual melting takes place. These summertime decreases in runoff can be related to cooling (Forsythe *et al.*, 2017; Fowler and Archer, 2006; Tahir *et al.*, 2011). While the temperature records indicate an overall warming trend over the region, especially during the spring (Asad *et al.*, 2016; Fowler and Archer, 2006; Hasson *et al.*, 2015; Iqbal *et al.*, 2016), the direction of the mean temperature trend in summer is negative and coexisting with summertime decrease in runoff (Forsythe *et al.*, 2017; Fowler and Archer, 2006).

3.2 HUMIDITY AND CLOUD COVER

The in-situ meteorological observations indicate an increasing trend in regional humidity for all seasons of the year (Figures A. 2g & h, Tables S13 & S14), with highest rates in the summer i.e., 0.55 ± 0.12 & 0.82 ± 0.33 hPa/10yr in the morning (08-AM) and the afternoon (05-PM), respectively, and lowest in the winter (0.08 ± 0.04 & 0.13 ± 0.1 hPa/10yr). The summer season is also marked by a positive trend in cloud cover (0.14 & 0.17 Okta/10yr, see Figures A.2i & j, Tables S15 & S16). Increased cloudiness affects the snow and glaciated ice in several different ways. First, cloudiness is concurrent with precipitation during the summers (0.07 ± 0.03 mmD-1/10yr, see Table S23). Increased afternoon cloudiness (compared to morning) throughout the year is indicative of increased convective uplift at the regional scale, and indicates enhanced micro and mesoscale convective precipitation (especially in the summer). Second, at very high elevations, radiation is responsible for 80-85% of the ablation (Hewitt, 2005), while absorption and reflection back to space of the incoming solar radiation by clouds result in less melt occurring. Third, increased daytime cloudiness can decrease the underlying temperatures, and help to reduce the melting of snow and glaciated ice that would be caused by exposure to warm winds. Finally, decreased insolation can affect regional hydro-climate through its highly correlated change in wind speed (Xu *et al.*, 2006).

The observed increases in summer season atmospheric water vapor (WV) can be related to the possible penetration of monsoon moisture from the Arabian Sea into the western HKH region. Using the relationship between relative depletion in the heavy oxygen-isotope of

precipitated water and its proximity from the moisture source, Wake [1989] found that the summer monsoon from the Arabian Sea contributes nearly one-third of the annual precipitation in the Karakoram region. He also found that higher concentrations of $Na^{+}$and Cl^{-} are related to summer precipitation and that the mean $Na^{+}:Cl^{-}$ ratio of Karakoram is 0.54, which is very close to the ratio of sea-water (i.e., 0.56). Furthermore, observational analysis and all forcing (natural solar variations, volcanoes, anthropogenic well-mixed greenhouse gases, ozone, aerosols, and land use) ensemble simulations of June-September average precipitation for the period of 1959-1999 presented by Bollasina et al. [2011] indicate the emergence of a coherent large-scale pattern, in which drying over central-northern India coinciding with slight wetting over Pakistan is prominent. Therefore, the increase in summer season precipitation observed in the study areas (Figure A. 2p, Table S23) is more related to monsoon intrusion from the Arabian Sea than to western disturbances, especially when the mid-latitude jet stream (Archer and Caldeira, 2008) and monsoon low level jet (LLJ) are moving poleward with the widening of the tropical belt (Sandeep and Ajayamohan, 2015). Recent research (Forsythe *et al.*, 2017) has related regional atmospheric circulation to Karakoram temperatures and glacial melt through their Karakoram Zonal Index, and has suggested plausible weakening of the summer monsoon. However, this is not the case, because the index they have defined is indifferent to the south westerly monsoon from the Arabian Sea, which is the main source of summer precipitation in the Karakoram region (Wake, 1989).

3.3 WIND SPEED

Increases in WV have profound effects on surface net radiation flux and its partitioning by limiting evapotranspiration, especially when the wind speed is not strong enough to effectively remove the water vapor from a specific location. In the context of the study area, near surface wind speeds have not received much attention. Along with vapor pressure and radiation, wind speed is an important meteorological variable that affects the partitioning of available energy into sensible and latent heat fluxes (Singh and Bengtsson, 2005; Winiger *et al.*, 2005). The data indicate that near-surface wind speeds throughout the region have been decreasing over time

(year around) (Figures A. 2k & l, Tables S17 & S18). This is especially true in summer, with larger decreases in the afternoon (-0.17 ± 0.16 ms-1/10yr) in comparison to the morning (-0.08 ± 0.09 ms-1/10yr). The change in the morning wind speed is directly related to the minimum temperature (with their statistically significant linear correlation of 0.47). Decreased wind speeds can reduce the thickness of the boundary layer, and reduce the mechanical mixing of sensible heat, leading to reduced temperatures in the lowest part of the troposphere. The inversion signature is dominant in Kabul radiosonde reports as presented in Figure S3 and S4 (see S.I for discussion on temperature inversion).

The weakening in wind speed observed in the valley floors is relevant to the elevated glacier regions, as well. For instance, under melting conditions when the glacier surface temperature is almost 0°C, the temperature gradient of the glacier surface with ambient air temperature regulates katabatic flows (Oerlemans and Grisogono, 2002). Furthermore, van den Broeke [1997] mentioned that during periods with overcast skies and low temperature, a low potential temperature difference at different locations of a glacier is observed, indicating a weak or absent katabatic wind system. Considering these, katabatic flows are expected to decrease with a decrease in summer season air temperature and increase in cloudiness, which will further reduce melting on the glacier surface by reduction in adiabatic compression.

3.4 NET RADIATION

Note that incoming insolation at the surface decreases as cloudiness increases, and the observations show a decrease in net radiation at the regional scale, especially in the summer season (-0.12 ± 0.02 MJm2D-1/10yr; Figure A. 2q, Table S21). Moreover, in general, net radiation decreases with elevation due to the increase in surface albedo and the decrease in incoming longwave radiation, whereas, the outgoing longwave radiation is more or less fixed by the melting ice/snow surface (van den Broeke, 1997).

3.5 POTENTIAL EVAPOTRANSPIRATION AND CLIMATE MOISTURE INDEX (CMI)

Decreases in the summer daytime temperature, wind speed, net radiation combined with an increase in cloudiness and WV favor reductions in potential evapotranspiration (ETp) that may further influence the aridity of the region. Therefore, as a next step, the meteorological observations are used to obtain estimates of evaporation and the CMI (see S.I. for details). As expected, ETp is decreasing throughout the year, and particularly during the summer (-0.24 ± 0.13 mmD-1/10yr; Figure A. 2o, Table S22). Note that the little decrease in spring season ETp (-0.08 ± 0.07 mmD-1/10yr) is credited mainly to a decrease in wind speed, whereas net radiation in spring season is increasing due to decrease in cloud cover (0.02 ± 0.05 MJm2D-1/10yr). Moreover, since summer is the main season in which snowmelt and ablation occur from the glaciers in the Karakoram and adjoining region, therefore decreases in summer available energy and ETp have a serious impact on the melting process.

Since the water and energy budgets of a watershed are inextricably intertwined, increases in the energy inputs to a watershed can lead to decreases in moisture, and vice versa. Data regarding ETp and precipitation can be used to estimate values for the CMI (Willmott and Feddema, 1992). Our estimates of the CMI indicate that, due to increased accumulated summer precipitation and decreased ETp, the region is acquiring a moisture surplus while becoming energy deficient in the summer, and at the overall annual scale (Figure A. 2r), thereby influencing the hydrological balance of the region.

4. DISCUSSION

In the ablation zone of the glacier, reductions in temperature (Fowler and Archer, 2006) and increases in moisture and cloudiness, when coupled with the reductions in insolation, can have a cooling effect, with decreased river runoff being the natural consequence (Archer and Fowler, 2004; Forsythe *et al.*, 2017; Fowler and Archer, 2006; Tahir *et al.*, 2011). On the other hand, in the accumulation zone of the glacier, increased cloudiness and precipitation (Archer and

Fowler, 2004; Nepal and Shrestha, 2015), thermal inputs through the release of latent heat (Quincey *et al.*, 2011), and increased temperatures of the summer snowfall, can act to facilitate the transfer of ice to lower parts of the glacier through surges or avalanches (Bolch *et al.*, 2012; Copland *et al.*, 2011; Quincey *et al.*, 2011, 2015; Rankl *et al.*, 2014). The combined effect can be to alter the mass balance of the glacier in a positive direction (Hewitt, 2005, 2007). The observational records presented here indicate that there is a coherent interplay of meteorological variables at the regional scale, which combines with non-climatic factors such as extreme vertical topographic gradients (Hewitt, 2005), heavy debris cover (Smiraglia *et al.*, 2007), steepness and thermal characteristics (Hewitt, 2005, 2007; Kapnick *et al.*, 2014; Quincey *et al.*, 2011; Rankl *et al.*, 2014) of the ice mass, to contribute to glacier growth and/or surges. While valley floor observations are sometimes considered to be insufficient in terms of longevity, suffering from statistically insignificant trends, and presenting contradictory signals due to local variations in elevation, slope and aspect (Bhutiyani *et al.*, 2007; Fowler and Archer, 2006; Kapnick *et al.*, 2014), the instrumental records available for the eastern Hindukush, western Karakoram and northwestern Himalaya centered in northern Pakistan (presented here) show an integrated signal with all locations indicating similar trends in hydro-climatic change at the regional scale. These observational meteorological records indicate that energy, mass and momentum fluxes have been acting to offset each other in increasing and decreasing patterns over the last five decades. Increases in water vapor, cloudiness and precipitation, and decreases in net radiation, near surface wind speed, and potential evapotranspiration have resulted in a positive hydrologic mass balance at the annual scale.

However, the results presented here show considerable variability. In summer, the combination of increased precipitation, decreased discharge, and diminished ETp results in a positive mass balance stored as glaciated ice. While for spring, the combination of decreased precipitation, increased runoff, and relatively little change in ETp indicates a negative mass balance. Given that spring season flows in the Indus river are mainly due to snow melt and for summer season contributions are mainly due to glacier melt, it can be assumed that the snowfields

are more vulnerable to early melt in the spring, whereas the glaciers are more protected from melting in the summer, especially when monsoon moisture is helping to nourish the glacier tops. Of course, given limited observational data, hydro-meteorological variables cannot be computed without employing some standard assumptions. Here, a fixed value of albedo has been assumed based on a review of the literature (Ming *et al.*, 2015). While changing albedo values can alter the absolute magnitudes of the different variables, they will not influence the seasonal trends of the computed variables provided albedo remains fixed over an entire season. On the other hand, extensive cloud cover that covers the whole horizon can have a dominant influence on incoming solar insolation, and so increased cloudiness can have a profound impact on ablation through obstruction of insolation, and by weakened katabatic winds on glacier surface that diminishes melting. Furthermore, increase in relative humidity (decrease in vapor pressure deficit) observed at the valley floors is related to its increase in the glaciated regions (high elevations) as vapor pressure deficit decreases with drop in temperature at elevation (Busing *et al.*, 1950), and therefore discourage sublimation on the high elevated glaciers. Collectively, all these assumptions are realistic and help to explain the contributions of climatic forcing in the establishment of the Karakoram Anomaly.

5. CONCLUSION

The analysis of the long-term meteorological dataset presented here shows a clear signal that the HKH watersheds have been moisture surplus and energy deficient in recent years. Further, nourishment of the glaciers in both winter and summer has been increasing lately. Overall, these results clearly support the existence of the “Karakoram Anomaly” due to a positive mass balance of glaciers in the region. Together with non-climatic factors this positive mass balance is facilitating the observed glacier surges.

ACKNOWLEDGMENTS, SAMPLES, AND DATA

We thank Pakistan Meteorological Department (PMD), Climate Data Processing Centre (CDPC, http://www.pmd.gov.pk/cdpc/home.htm), Water and Power Development Authority

(WAPDA, http://www.wapda.gov.pk/), and Dr. Ghulam Rasul (Director General, PMD) for providing the hydro-climatic data used in this study. The first author is grateful to the United States Education Foundation Pakistan (USEFP) for providing the Fulbright Scholarship that made this work possible. We acknowledge the climate data guide to distribute CRU data, the Global Modeling and Assimilation Office (GMAO) and the Goddard Earth Science Data Information Services Center (GES DISC) for the dissemination of MERRA, the European Centre for Medium-Range Weather Forecasts (ECMWF) for ERA-Interim, the National Centers for Environmental Information (NCEI) and Integrated Global Radiosonde Archive (IGRA) to distribute radiosonde data, the Global Land Cover Facility (GLCF) to distribute digital elevation model (DEM), and the global administrative areas (GADM) to distribute world's administrative boundaries. This study was funded by NASA (NNX14AM02G).

REFERENCES

Adler RF, Huffman GJ, Chang A, Ferraro R, Xie P-P, Janowiak J, Rudolf B, Schneider U, Curtis S, Bolvin D, Gruber A, Susskind J, Arkin P, Nelkin E. 2003. The Version-2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979–Present). *Journal of Hydrometeorology* **4**(6): 1147–1167. DOI: 10.1175/1525-7541(2003)004<1147:TVGPCP>2.0.CO;2.

Adler RF, Negri AJ. 1988. A Satellite Infrared Technique to Estimate Tropical Convective and Stratiform Rainfall. *Journal of Applied Meteorology*, 30–51. DOI: 10.1175/1520-0450(1988)027<0030:ASITTE>2.0.CO;2.

Arkin PA, Joyce R, Janowiak JE. 2017. The estimation of global monthly mean rainfall using infrared satellite data : The GOES precipitation index (GPI) The Estimation of Global Monthly Mean Rainfall Using Infrared Satellite Data : The GOES Precipitation Index (GPI). **7257**(July). DOI: 10.1080/02757259409532261.

Arkin P, Meisner B. 1987. The relationship between large-scale convective rainfall and cold cloud over the western hemisphere during 1982-84. *Monthly Weather Review*, 51–74. DOI: 10.1175/1520-0493(1987)115<0051:TRBLSC>2.0.CO;2.

Ashouri H, Hsu KL, Sorooshian S, Braithwaite DK, Knapp KR, Cecil LD, Nelson BR, Prat OP. 2015. PERSIANN-CDR: Daily precipitation climate data record from multisatellite observations for hydrological and climate studies. *Bulletin of the American Meteorological Society* **96**(1): 69–83. DOI: 10.1175/BAMS-D-13-00068.1.

Barrett EC. 2000. Satellite remote sensing of precipitation : progress and problems. *Remote Sensing Hydrology* **2000**(267): 3–10.

Becker A, Finger P, Meyer-Christoffer A, Rudolf B, Schamm K, Schneider U, Ziese M. 2013. A description of the global land-surface precipitation data products of the Global Precipitation Climatology Centre with sample applications including centennial (trend) analysis from 1901-present. *Earth System Science Data* **5**(1): 71–99. DOI: 10.5194/essd-5-71-2013.

Behrangi A, Imam B, Hsu K, Sorooshian S, Bellerby TJ, Huffman GJ. 2010. REFAME: Rain Estimation Using Forward-Adjusted Advection of Microwave Estimates. *Journal of Hydrometeorology* **11**: 1305–1321. DOI: 10.1175/2010JHM1248.1.

Berne A, Delrieu G, Creutin JD, Obled C. 2004. Temporal and spatial resolution of rainfall measurements required for urban hydrology. *Journal of Hydrology* **299**(3–4): 166–179. DOI: 10.1016/j.jhydrol.2004.08.002.

Chang CP, Wang Z, Ju J, Li T. 2004. On the relationship between western maritime continent monsoon rainfall and ENSO during northern winter. *Journal of Climate* **17**(3): 665–672. DOI: 10.1175/1520-0442(2004)017<0665:OTRBWM>2.0.CO;2.

Chen M, Shi W, Xie P, Silva VBS, Kousky VE, Higgins RW, Janowiak JE. 2008. Assessing objective techniques for gauge-based analyses of global daily precipitation. *Journal of Geophysical Research Atmospheres* **113**(4): 1–13. DOI: 10.1029/2007JD009132.

Chen M, Xie P, Janowiak JE, Arkin PA. 2002. Global Land Precipitation: A 50-yr Monthly Analysis Based on Gauge Observations. *Journal of Hydrometeorology* **3**(3): 249–266. DOI: 10.1175/1525-7541(2002)003<0249:GLPAYM>2.0.CO;2.

Daly C, Halbleib M, Smith JI, Gibson WP, Doggett MK, Taylor GH, Curtis J, Pasteris PP. 2008. Physiographically sensitive mapping of climatological temperature and precipitation

across the conterminous United States. *International Journal of Climatology* **28**(15): 2031–2064. DOI: 10.1002/joc.1688.

Dee DP, Uppala SM, Simmons AJ, Berrisford P, Poli P, Kobayashi S, Andrae U, Balmaseda MA, Balsamo G, Bauer P, Bechtold P, Beljaars ACM, van de Berg L, Bidlot J, Bormann N, Delsol C, Dragani R, Fuentes M, Geer AJ, Haimberger L, Healy SB, Hersbach H, Holm E V., Isaksen L, Kallberg P, Kohler M, Matricardi M, Mcnally AP, Monge-Sanz BM, Morcrette JJ, Park BK, Peubey C, de Rosnay P, Tavolato C, Thepaut JN, Vitart F. 2011. The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society* **137**(656): 553–597. DOI: 10.1002/qj.828.

Duchon CE, Biddle CJ. 2010. Undercatch of tipping-bucket gauges in high rain rate events. *Adv. Geosci* **25**: 11–15. DOI: 10.5194/adgeo-25-11-2010.

Duchon CE, Essenberg GR. 2001. Comparative rainfall observations from pit and aboveground rain gauges with and without wind shields. *Water Resources Research* **37**(12): 3253–3263. DOI: 10.1029/2001WR000541.

Harris I, Jones PD, Osborn TJ, Lister DH. 2014. Updated high-resolution grids of monthly climatic observations - the CRU TS3.10 Dataset. *International Journal of Climatology* **34**(3): 623–642. DOI: 10.1002/joc.3711.

Haylock MR, Hofstra N, Klein Tank AMG, Klok EJ, Jones PD, New M. 2008. A European daily high-resolution gridded data set of surface temperature and precipitation for 1950-2006. *Journal of Geophysical Research Atmospheres* **113**(20). DOI: 10.1029/2008JD010201.

Hou AY, Kakar RK, Neeck S, Azarbarzin AA, Kummerow CD, Kojima M, Oki R, Nakamura K, Iguchi T. 2014. The global precipitation measurement mission. *Bulletin of the American Meteorological Society* **95**(5): 701–722. DOI: 10.1175/BAMS-D-13-00164.1.

Huffman GJ, Adler RF, Morrissey MM, Bolvin DT, Curtis S, Joyce R, McGavock B, Susskind J. 2001. Global Precipitation at One-Degree Daily Resolution from Multisatellite Observations. *Journal of Hydrometeorology* **2**(1): 36–50. DOI: 10.1175/1525-7541(2001)002<0036:GPAODD>2.0.CO;2.

Huffman GJ, Bolvin DT, Braithwaite D, Hsu K, Joyce R, Xie P, Yoo S-H. 2014. NASA global precipitation measurement (GPM) integrated multi-satellite retrievals for GPM (IMERG). *Algorithm Theoretical Basis Document (ATBD), NASA/GSFC, Greenbelt, MD, USA*

Huffman GJ, Bolvin DT, Nelkin EJ, Wolff DB, Adler RF, Gu G, Hong Y, Bowman KP, Stocker EF. 2007. The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales. *Journal of Hydrometeorology* **8**(1): 38–55. DOI: 10.1175/JHM560.1.

Joyce RJ, Janowiak JE, Arkin PA, Xie P. 2004. CMORPH: A Method that Produces Global Precipitation Estimates from Passive Microwave and Infrared Data at High Spatial and Temporal Resolution. *Journal of Hydrometeorology* **5**(3): 487–503. DOI: 10.1175/1525-7541(2004)005<0487:CAMTPG>2.0.CO;2.

Kidd C, Huffman G. 2011. Global precipitation measurement. *Meteorological Applications* **18**(3): 334–353. DOI: 10.1002/met.284.

Kidd C, Kniveton DR, Todd MC, Bellerby TJ, Kidd C, Kniveton DR, Todd MC, Bellerby TJ. 2003. Satellite Rainfall Estimation Using Combined Passive Microwave and Infrared Algorithms. *Journal of Hydrometeorology* **4**(6): 1088–1104. DOI: 10.1175/1525-7541(2003)004<1088:SREUCP>2.0.CO;2.

Kubota T, Hashizume H, Shige S, Okamoto K, Aonashi K, Takahashi N, Ushio T, Kachi M. 2006. Global precipitation map using satelliteborne microwave radiometers by the GSMaP project: Production and validation. *International Geoscience and Remote Sensing Symposium (IGARSS)*, 2584–2587. DOI: 10.1109/IGARSS.2006.668.

Kummerow C, Barnes W, Kozu T, Shiue J, Simpson J. 1998. The Tropical Rainfall Measuring Mission (TRMM) sensor package. *Journal of Atmospheric and Oceanic Technology* **15**(3): 809–817. DOI: 10.1016/0273-1177(94)90210-0.

Lau KM, Wu HT. 2001. Principal modes of rainfall-SST variability of the Asian Summer Monsoon: A reassessment of the Monsoon-ENSO relationship. *Journal of Climate* **14**(13): 2880–2895. DOI: 10.1175/1520-0442(2001)014<2880:PMORSV>2.0.CO;2.

Mariotti A. 2007. How ENSO impacts precipitation in southwest central Asia. *Geophysical Research Letters* **34**(16): 2–6. DOI: 10.1029/2007GL030078.

Nešpor V, Krajewski WF, Kruger A. 2000. Wind-Induced Error of Raindrop Size Distribution Measurement Using a Two-Dimensional Video Disdrometer. *Journal of Atmospheric and Oceanic Technology* **17**(11): 1483–1492. DOI: 10.1175/1520-0426(2000)017<1483:WIEORS>2.0.CO;2.

Osborn TJ, Hulme M. 1998. Evaluation of the European daily precipitation characteristics from the atmospheric model intercomparison project. *International Journal of Climatology* **18**(5): 505–522. DOI: 10.1002/(SICI)1097-0088(199804)18:5<505::AID-JOC263>3.0.CO;2-7.

Reinert D, Prill F, Frank H, Zängl G. 2016. A global satellite-assisted precipitation climatology. *Deutscher Wetterdienst, Research and Development (FE1 3)* (1994): 2000. DOI: 10.5676/DWD.

Schneider U, Fuchs T, Meyer-Christoffer A, Rudolf B. 2011. Global Precipitation Analysis Products of the GPCC. Global Precipitation Climatology Centre (GPCC). *Gpcc.Dwd.De* 12.

Scofield RA, Kuligowski RJ. 2003. Status and Outlook of Operational Satellite Precipitation Algorithms for Extreme-Precipitation Events. *Weather and Forecasting* **18**(6): 1037–1051. DOI: 10.1175/1520-0434(2003)018<1037:SAOOOS>2.0.CO;2.

Shepard D. 1968. A two-dimensional interpolation function for irregularly-spaced data. *Proceedings of the 1968 23rd ACM national conference on* - 517–524. DOI: 10.1145/800186.810616.

Sorooshian S, Hsu KL, Gao X, Gupta H V., Imam B, Braithwaite D. 2000. Evaluation of PERSIANN system satellite-based estimates of tropical rainfall. *Bulletin of the American Meteorological Society* **81**(9): 2035–2046. DOI: 10.1175/1520-0477(2000)081<2035:EOPSSE>2.3.CO;2.

Stahl K, Moore RD, Floyer JA, Asplin MG, McKendry IG. 2006. Comparison of approaches for spatial interpolation of daily air temperature in a large region with complex topography

and highly variable station density. *Agricultural and Forest Meteorology* **139**(3–4): 224–236. DOI: 10.1016/j.agrformet.2006.07.004.

Stephens GL, Vane DG, Boain RJ, Mace GG, Sassen K, Wang Z, Illingworth AJ, O'Connor EJ, Rossow WB, Durden SL, Miller SD, Austin RT, Benedetti A, Mitrescu C. 2002. The cloudsat mission and the A-Train: A new dimension of space-based observations of clouds and precipitation. *Bulletin of the American Meteorological Society* **83**(12): 1771–1790+1742. DOI: 10.1175/BAMS-83-12-1771.

Tapiador FJ, Navarro A, Levizzani V, García-Ortega E, Huffman GJ, Kidd C, Kucera PA, Kummerow CD, Masunaga H, Petersen WA, Roca R, Sánchez J-L, Tao W-K, Turk FJ. 2017. Global precipitation measurements for validating climate models. *Atmospheric Research* **197**(May): 1–20. DOI: 10.1016/j.atmosres.2017.06.021.

Tapiador FJ, Turk FJ, Petersen W, Hou AY, García-Ortega E, Machado LAT, Angelis CF, Salio P, Kidd C, Huffman GJ, de Castro M. 2012. Global precipitation measurement: Methods, datasets and applications. *Atmospheric Research*. Elsevier B.V. **104**–**105**: 70–97. DOI: 10.1016/j.atmosres.2011.10.021.

Taylor KE. 2001. in a Single Diagram. *Journal of Geophysical Research* **106**(D7): 7183–7192. DOI: 10.1029/2000JD900719.

Teegavarapu RS V, Chandramouli V. 2005. Improved weighting methods, deterministic and stochastic data-driven models for estimation of missing precipitation records. *Journal of Hydrology* **312**(1–4): 191–206. DOI: 10.1016/j.jhydrol.2005.02.015.

Turk F. J, Hawkins J, Smith EA, Marzano FS, Mugnai A, Levizzani V. 2000. Combining SSM/I, TRMM and Infrared Geostationary Satellite Data in a Near-realtime Fashion for Rapid Precipitation Updates: Advantages and Limitations. *The 2000 EUMETSAT Meteorological Satellite Data Users' Conference* 452–459.

Vicente GA, Scofield RA, Menzel WP. 1998. The 0perational GOES infrared rainfall estimation technique. *Bulletin of the American Meteorological Society* **79**(9): 1883–1898. DOI: 10.1175/1520-0477(1998)079<1883:TOGIRE>2.0.CO;2.

Viney NR, Bates BC. 2004. It never rains on Sunday: the prevalence and implications of untagged multi-day rainfall accumulations in the Australian high quality data set. *International Journal of Climatology* **24**(9): 1171--1192.

Willmott CJ, Rowe CM, Philpot WD. 1985. Small-Scale Climate Maps: A Sensitivity Analysis of Some Common Assumptions Associated with Grid-Point Interpolation and Contouring. *Cartography and Geographic Information Science* **12**(1): 5–16. DOI: 10.1559/152304085783914686.

Xie P, Arkin PA. 1997. Global Precipitation: A 17-Year Monthly Analysis Based on Gauge Observations, Satellite Estimates, and Numerical Model Outputs. *Bulletin of the American Meteorological Society* **78**(11): 2539–2558. DOI: 10.1175/1520-0477(1997)078<2539:GPAYMA>2.0.CO;2.

Xie P, Chen M, Yang S, Yatagai A, Hayasaka T, Fukushima Y, Liu C. 2007. A Gauge-Based Analysis of Daily Precipitation over East Asia. *Journal of Hydrometeorology* **8**(3): 607–626. DOI: 10.1175/JHM583.1.

Xie P, Janowiak JE, Arkin PA, Adler R, Gruber A, Ferraro R, Huffman GJ, Curtis S. 2003. GPCP pentad precipitation analyses: An experimental dataset based on gauge observations and satellite estimates. *Journal of Climate* **16**(13): 2197–2214. DOI: 10.1175/2769.1.

Yatagai A, Kamiguchi K, Arakawa O, Hamada A, Yasutomi N, Kitoh A. 2012. Aphrodite constructing a long-term daily gridded precipitation dataset for Asia based on a dense network of rain gauges. *Bulletin of the American Meteorological Society* **93**(9): 1401–1415. DOI: 10.1175/BAMS-D-11-00122.1.

Table A. 1 a) List of meteorological sites used in the study with their latitude, longitude, elevation (meters) and description of the data used in the study. Symbol of ☀ indicates that monthly mean of maximum, mean and minimum temperature from that site is utilized and symbol of ☁ indicates that additional data such as monthly means of synoptic weather data reported by meteorological stations in morning (08:00 a.m., local time) and afternoon (05:00 p.m., local time) along with accumulated precipitation are utilized for the study. b) List of observed and derived meteorological variables utilized to perform analyses.

a

Station Name	Latitude	Longitude	Elevation (m)	Data Type
Astor	35.3	74.9	2167	☀/☁
Bunji	35.6	74.6	1372	☀/☁
Chillas	35.6	74.1	1250	☀/☁
Chitral	35.8	71.8	1499	☀
Dir	35.2	71.8	1375	☀
Drosh	35.5	71.7	1463	☀
Gilgit	35.9	74.3	1459	☀/☁
Gupis	36.2	73.4	2156	☀/☁
Garhi Dupatta	34.1	73.6	812	☀
Kotli	33.5	73.9	613	☀
Muzaffarabad	34.2	74	701	☀
Skardu	35.3	75.6	2317	☀/☁

b

Variable Name	Units	Retrieval Method
Maximum Temperature	℃	Observed
Minimum Temperature	℃	Observed
Mean Temperature	℃	Derived
Diurnal Temperature Range	℃	Derived
Precipitation	mm	Observed
Dry Bulb Temperature	℃	Observed
Wind Speed	ms^{-1}	Observed
Cloud Cover Fraction	Okta	Observed
Surface Pressure	kPa	Observed
Relative Humidity	%	Observed
Vapor Pressure	hPa	Derived
Net Radiation	$MJm^{-2}d^{-1}$	Derived
Potential Evapotranspiration	mm	Derived
Climatic Moisture Index	%	Derived

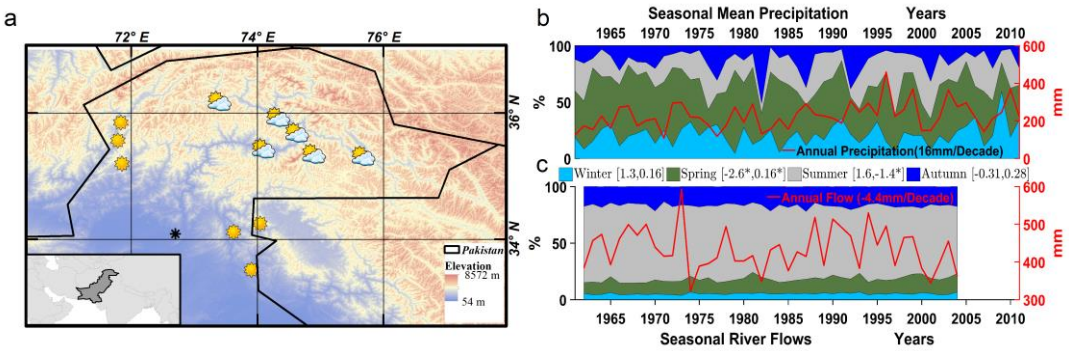


Figure A. 1. Study Area and Hydrological Budget. a) Map showing the location of HKH centered in Northern Pakistan, and the meteorological sites used in this study. The symbols of (☀) and (🌅) are explained in Table 1 and (∗) indicates Tarbela Dam's location. Elevation from mean sea level is presented in legend. Right panels show the accumulated percentage (shaded, left axis) of seasonal contributions in mean regional annual precipitation (1961-2011), b), and discharge recorded at Tarbela Dam (April 1961-March 2005), c), while red line in (b) indicates regional mean of annual accumulated precipitation (mm, right axis) and in (c) it indicates the total annual sum of specific runoff (mm) at Tarbela Dam. The legend brackets indicate trends in seasonal contributions (from winter to autumn, in % per decade) to the mean regional precipitation (b) and discharge (c), respectively. Statistical significance (95% confidence interval) of the trends is marked with an asterisk sign.

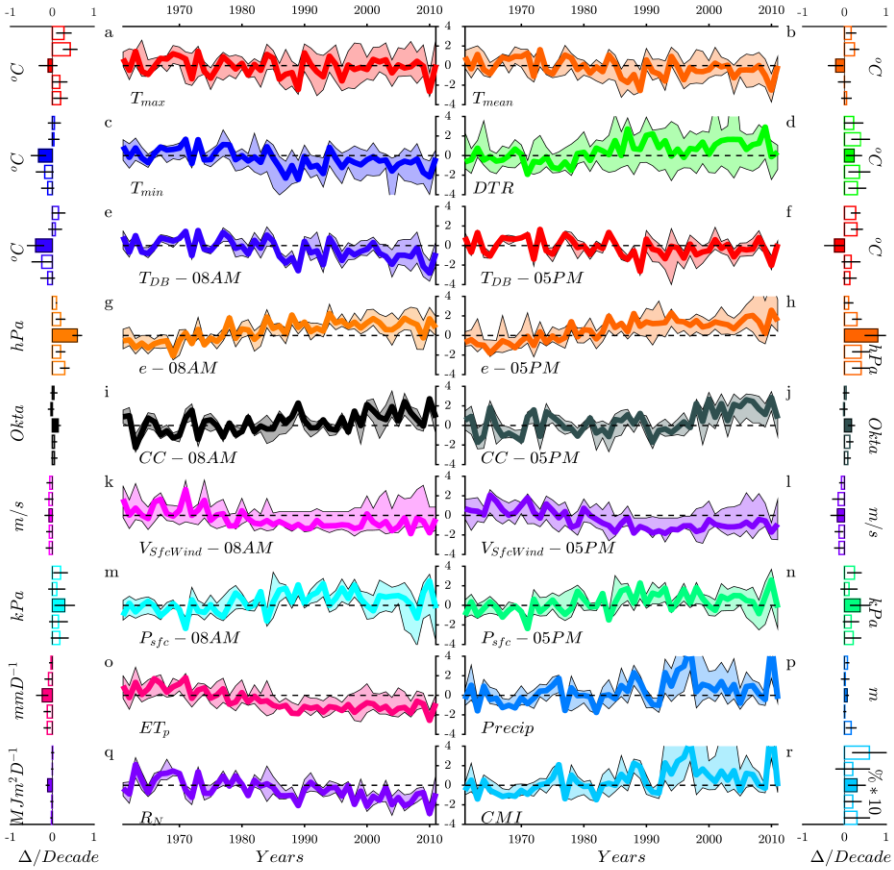


Figure A. 2. Bar plots (on both sides) indicate mean regional trends of the seasonal means of different meteorological variables with one standard deviation on both sides of the bar indicated by the horizontal line; from top to bottom are winter, spring, summer, autumn and annual, respectively. Time series plots show the corresponding summer anomalies standardized with respect to mean and standard deviation for 1961-1990. This base period is selected to present change in variables with respect to Pre-Karakoram Anomaly conditions. a) maximum temperature; b) mean temperature; c) minimum temperature; d) diurnal temperature range; e) dry bulb temperature at 08:00 a.m.; f) dry bulb temperature at 05:00 p.m.; g) vapor pressure at 08:00 a.m.; h) vapor pressure at 05:00 p.m.; i) cloud cover at 08:00 a.m.; j) cloud cover at 05:00

p.m.; k) near surface wind speed at 08:00 a.m.; l) near surface wind speed at 05:00 p.m.; m) surface pressure at 08:00 a.m.; n) surface pressure at 05:00 p.m.; o) potential evapotranspiration; p) accumulated precipitation; q) net radiation; r) climate moisture index.

SUPPORTING INFORMATION

A HYDRO-METEOROLOGICAL PERSPECTIVE ON THE KARAKORAM ANOMALY USING UNIQUE VALLEY-BASED SYNOPTIC WEATHER OBSERVATIONS

Text S1 to S3

Figures S1 to S5

Tables S1 to S25

Text S1

Introduction

The following material provides additional information on study area, location of observatories, data used in the study, description of variables, ratio of missing values, homogeneity of time series, and methodology to compute available energy, potential evapotranspiration and Climate Moisture Index (CMI). Further, it provides discussion on temperature inversion reported by radiosonde data from Kabul, Afghanistan and its relation to hydro-climatic signature in the region in Text S2. Text S3 compares in-situ observed variables with gridded observed and reanalysis datasets.

Study Area

The study area is comprised of eastern Hindukush, western Karakoram and northwestern Himalayan mountain ranges that are juxtaposed in the center of the province of Gilgit-Baltistan, situated in Northern Pakistan (see Figure S1). This area supports several concentrations of glaciers at high altitude. To analyze glaciers response to the prevailing climatic conditions, meteorological sites with long-term climatic records are selected. All such stations are operated by Pakistan Meteorological Department (PMD) and most of them are located on the valley floors. Latitude, longitude, and elevation of these sites are presented in Table 1.

Description of Variables

Monthly mean maximum and minimum temperatures, monthly accumulated precipitation, along with monthly mean synoptic weather observations of dry bulb temperature,

wind speed, cloud cover fraction and relative humidity observed at meteorological stations are used in this study. A list of observed and derived meteorological variables is presented in Table 2. Monthly mean maximum and minimum temperatures are available for 12 sites, whereas the monthly means of synoptic weather observations are available from 6 sites as presented in Table 1.

Monthly averages of synoptic weather observations, such as, dry bulb temperature, wind speed, cloud cover fraction, surface pressure, and relative humidity observed in the morning at 08-AM local time and 05-PM local time are used to compute vapor pressure, saturated vapor pressure, and potential evapotranspiration for each site (Shuttleworth, 2012). Regional averages at seasonal scales are computed for the seasonal time series analysis.

As evapotranspiration is strongly dependent on the net radiation, it is imperative to evaluate the net radiation available in the region. Without direct measurements of net radiation, we use an empirical method (Shuttleworth, 2012) to estimate the daily total solar radiation S reaching the surface:

$$S = [a_s + (1 - c)b_s]S_o \qquad \text{(S1)}$$

where c is the observed cloud cover fraction, and S_o is the solar constant (1367 wm-2) with adjustments using the eccentricity factor, latitude, sun hour angle, and solar declination for each Julian day at each location (Shuttleworth, 2012). Here, the 15th day of each month is taken to compute radiation budget on monthly scale. Empirical values of a_s and b_s can be derived locally by using measurements of insolation on overcast days to give a_s and on days with continuous bright sunshine to give (a_s + b_s). Typical values derived this way are a_s = 0.25 and b_s = 0.5 (Shuttleworth, 2012), which are used in this study. The surface net radiation is then estimated as

$$S_n = (1 - a)S \qquad \textit{(S2)}$$

where the albedo over HKH is taken as a constant [0.55 as suggested by Ming et al., 2015]. Similarly, we use a simple empirical method (Shuttleworth, 2012) to estimate the daily average net-longwave radiation:

$$L_n = -f\varepsilon'\sigma T_{air}^4 \qquad \text{(S3)}$$

where T_{air} is the daily average air temperature (i.e., average of morning and afternoon temperature), σ is Stephen-Boltzmann constant, and ε' is the surface effective emissivity:

$$\varepsilon' = 0.34 - 0.14\sqrt{e_d} \qquad \text{(S4)}$$

with e_d being the daily average vapor pressure computed using relative humidity and parametric form of Clausius-Clapeyron relationship that is based on air temperature. The factor f is an empirical cloud factor which, for arid conditions, can be calculated as:

$$f = 1.35\left(\frac{S}{S_{clear}}\right) - 0.35 \qquad \text{(S5)}$$

where S is computed from equation (S1) and S_{clear} is the value under clear-sky condition. Then the daily average net radiation flux R_n is simply the sum of S_n and L_n.
We also use the surface energy balance method (Allen *et al.*, 1998) to estimate potential evapotranspiration:

$$ET_p = \frac{R_n + \gamma\frac{900}{T+273}U_2(e_s - e_a)}{\Delta + \gamma(1 + 0.34U_2)} \qquad [mmd^{-1}] \qquad \text{(S6)}$$

where U_2 is observed wind speed at 2m height, T is the observed air temperature, e_a is actual vapor pressure, e_s is saturated vapor pressure, and psychrometric constant is defined as $\gamma = \frac{C_P P}{\varepsilon\lambda}$, where C_P is specific heat at constant pressure, ε is 0.622, P is atmospheric pressure, λ is latent heat of vaporization , and $\Delta = \frac{4098\, e_s}{(237.3+T)^2}$.

Finally, a dimensionless Climatic Moisture Index (CMI) (Willmott and Feddema, 1992) is adopted to investigate the state of the moisture in the HKH region:

$$\text{CMI} = 100\left[\frac{Pr}{ET_\text{p}} - 1\right] \qquad \text{(S7)}$$

where Pr is precipitation, and ET_p is computed from equation (S6). The CMI value is constrained between -100% and +100% with wet climate showing a positive CMI and dry climate a negative CMI.

Missing Values

Ratio of missing values to the length of the time series of all variables at seasonal time series are reported in Tables S1 and S2.

Homogeneity Test

(Archer and Fowler, 2004; Fowler and Archer, 2006) reported that the PMD and WAPDA follow standard meteorological measurement practices established by Indian Meteorological Department (IMD) in 1891. Using double mass curve approach they have found inhomogeneities in the winter minimum temperature around 1977 only at Bunji station among four low altitude stations analyzed. Forsythe et al., [2014] reported the homogeneity of Gilgit, Skardu and Astore stations for annual mean temperature during the period 1961–1990 while, Del Río et al., [2013] reported the homogeneity for the temperature record from the Gilgit, Gupis, Chillas, Astore and Skardu stations during 1952–2009. Hasson et al., [2015] reported inhomogeneity in minimum temperature but they never corrected it and emphasized on careful interpretation of results based on time series.

Here, homogeneity of all variables is tested at annual scale using Standard Normal Homogeneity Test (SNHT) as proposed by (Alexandersson, 1986; Tuomenvirta, 2002). Regional mean time series is adopted as a reference to check inhomogeneities in each candidate site. Years of major inhomogeneities of each variable, tested at 99% significance level, are reported in Tables S3 to S5. Not a single station reported a shift, that can be attributed to change in instrument location or observation practices. However, multiple shifts in the mean of time series at its different sections indicate strong environmental influences. Therefore, SNHT is again applied on the detrended time series and the years of major inhomogeneities are reported in Table S6. It is concluded that inhomogeneities reported by these time series are strongly influenced by the external factors and such shifts in mean of the time series at different sections cannot be adjusted by using a reference time series of another station that is dominated by the influence of different elevation, slope and aspect.

Trends and Significance

Seasonal trends of all the variables and their mean regional trends with standard deviation range are presented in Tables S7 to S23. Significance of the trends is estimated at 95% confidence interval and presented with bold fonts of reported trends.

Text S2

Temperature Inversion

Salient characteristics of reports from radiosonde data obtained from Kabul site (nearly 500 km from central Karakoram) are presented here. Firstly, in-situ observed data reports decrease (rather than increase) in minimum temperature along with increase in cloudiness in summer season, which is contrary to the usual perspective [e.g., Fowler and Archer, 2006]. However, the decrease in minimum temperature along with increasing cloud cover is plausible because of the observed decrease in near surface wind speed (08-AM, local time) that reduces the mechanical mixing of the lower troposphere. Since the lower troposphere is characterized by the temperature inversion at night (as found in radiosonde profiles of lower troposphere in summer), decreased mixing tends to reduce the minimum temperature. Secondly, as seen from Figure A. 2m & n of the main text, the surface pressure trend is positive throughout the year, while being a) strongest in the summer season and b) greater in the afternoon than in the morning. The regional warming that occurs at lower elevations is reflected in an increase in the surface pressures at elevated sites (Moore, 2012; Toumi *et al.*, 1999) and an increase in the 700 hPa geopotential height (Figure S2). This warming is also reflected in radiosonde data reports from Kabul, which shows positive lapse rate trends during all seasons. These trends are stronger at night (Figure S3) than during the daytime (Figure S4) in both the summer and autumn seasons, and vice versa in the winter season. Moreover, in recent years, the nighttime inversions are not as strong as they were previously. Temperature inversions in lower troposphere during the summer season are mainly due to the low temperatures occurring at sites around 2000 masl. While the inversion signature is still present in recent radiosonde profiles (Figure S3), however, it is not as strong as it was in 1980's. Radiosonde reports from Kabul are not continuous across the study time period. Therefore, only years with complete monthly reports are selected for the comparison.

Text S3

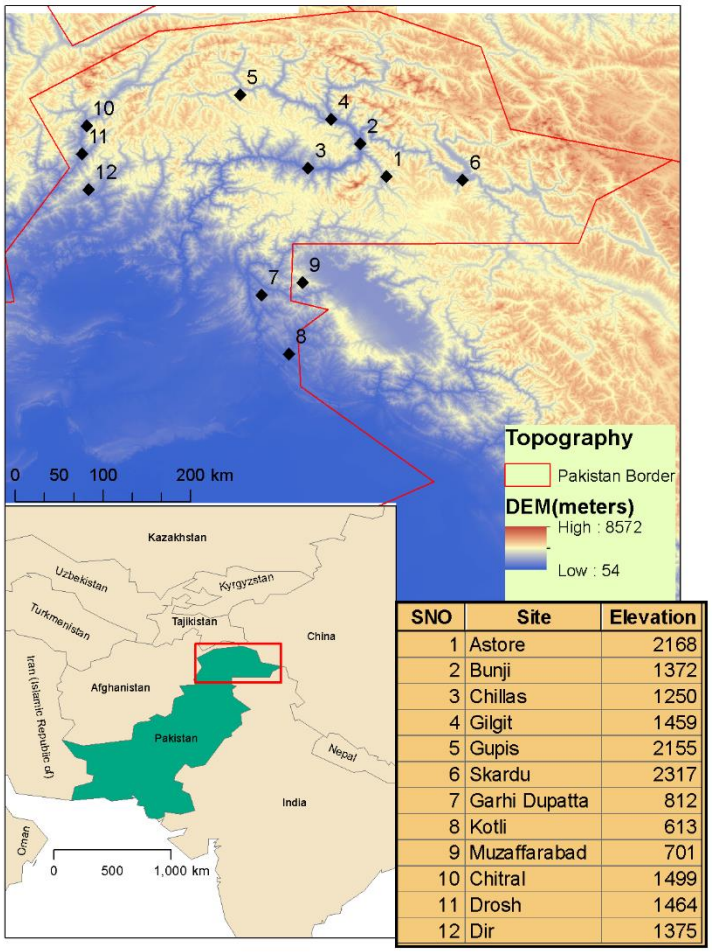
Comparison of Gridded Observed and Reanalysis

The role of the regional atmosphere in facilitation of the Karakoram anomaly has been suggested before (Hewitt, 2005). However, because long term meteorological observations have not previously been available, many previous studies have used global and regional climate models and gridded datasets to attempt to understand the meteorological characteristics facilitating this behavior (Ahmed *et al.*, 2011; Archer and Fowler, 2004; Asad *et al.*, 2016; Cheema and Bastiaanssen, 2012; Cook *et al.*, 2003; Kapnick *et al.*, 2014; Kumar *et al.*, 2015; Zafar *et al.*, 2016). These gridded observed and reanalysis data are spatially and temporally more continuous than instrumental records, and have therefore been adopted by scientific community as an alternate source of information. Figure S5 compares temperature trends from the gridded CRU TS 3.21 dataset (Harris *et al.*, 2014) and MERRA-2 reanalysis dataset (Rienecker *et al.*, 2011) for the summer season with the instrumental record used here. The negative trends in summer maximum temperatures reported by the instrumental record are indicated by the CRU dataset, as well. Compared to the CRU, the instrumental records in the Karakoram region indicate more cooling at higher altitude sites and greater warming at lower altitude sites. Similarly, all instrumental records report cooling in the mean and minimum temperatures, whereas CRU indicates significant warming throughout the region. Finally, the instrumental record indicates varying degrees of increase in diurnal temperature range (DTR) across the sites, whereas the CRU indicates a consistent decrease in DTR throughout the region from west to east.

Similarly, while the instrumental records show increases in maximum temperatures throughout the region (for the period of analysis), MERRA-2 indicates a decrease in maximum temperature in the western part centered in northern Pakistan and an increase in the eastern part. Meanwhile the instrumental record indicates a negative trend for minimum temperatures, whereas MERRA-2 indicates a positive trend. While the mean temperature trends reported by both instrumental record and MERRA-2 are more in agreement, they are not in agreement with the DTR trends (see Figure S5 for other variables).

This disagreement between the trends reported by instrumental observations and the gridded observed and the reanalysis data sets indicates that tremendous care should be taken

when using the latter, as they may not properly represent the changes in climate occurring in this region.



SNO	Site	Elevation
1	Astore	2168
2	Bunji	1372
3	Chillas	1250
4	Gilgit	1459
5	Gupis	2155
6	Skardu	2317
7	Garhi Dupatta	812
8	Kotli	613
9	Muzaffarabad	701
10	Chitral	1499
11	Drosh	1464
12	Dir	1375

Figure S1. Locations of Pakistan and study area in South Asia are presented. Subpanel shows a list of sites used in the study with their respective elevation in meters.

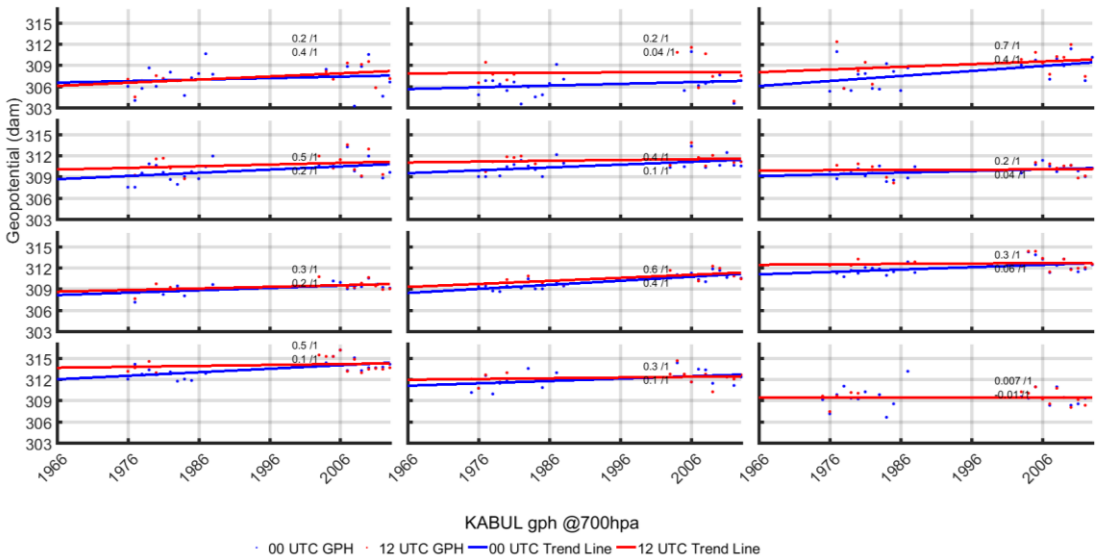


Figure S2. Monthly 700 hPa geopotential height trends (decameters/10yr) reported from radiosonde launched from Kabul, Afghanistan. Red and blue lines denote 700 hPa trend at 00 UTC and 12 UTC, respectively. Decadal trend is presented in numerator and significance is presented in denominator (1 = significant, 0 = non-significant at 95% confidence interval) where upper legend is referred to 00 UTC and lower is to 12 UTC.

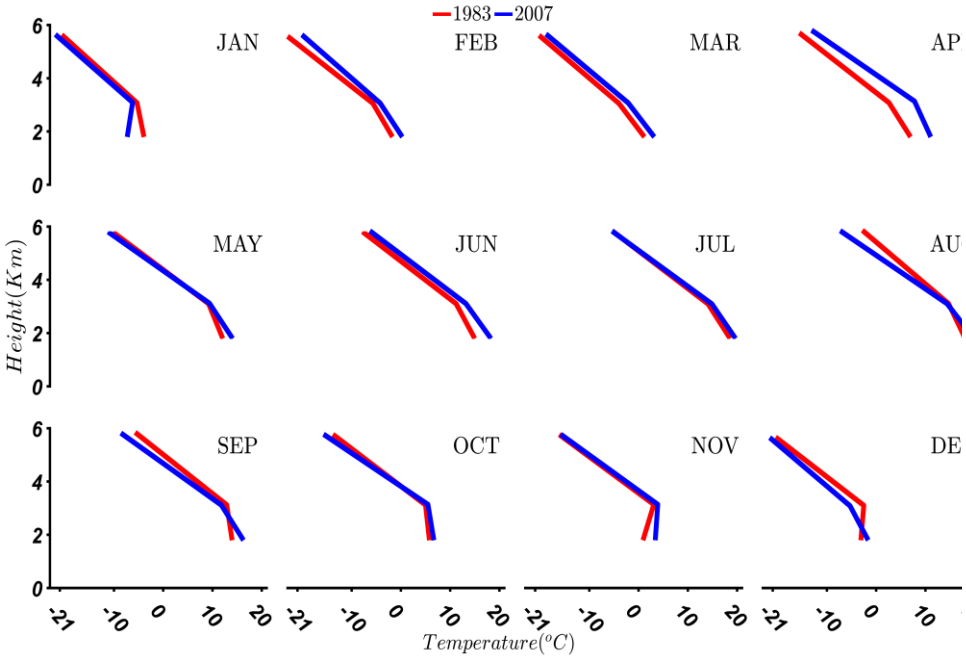


Figure S3. Monthly mean temperature from surface up to 6 Km reported by radiosonde launched at Kabul, Afghanistan in the Years 1980 and 2007 at 00-UTC (0430 AM, local time). These two years are selected as representatives of the decades of the 1980s and 2000s, respectively, as there are no missing values in those years at 00-UTC.

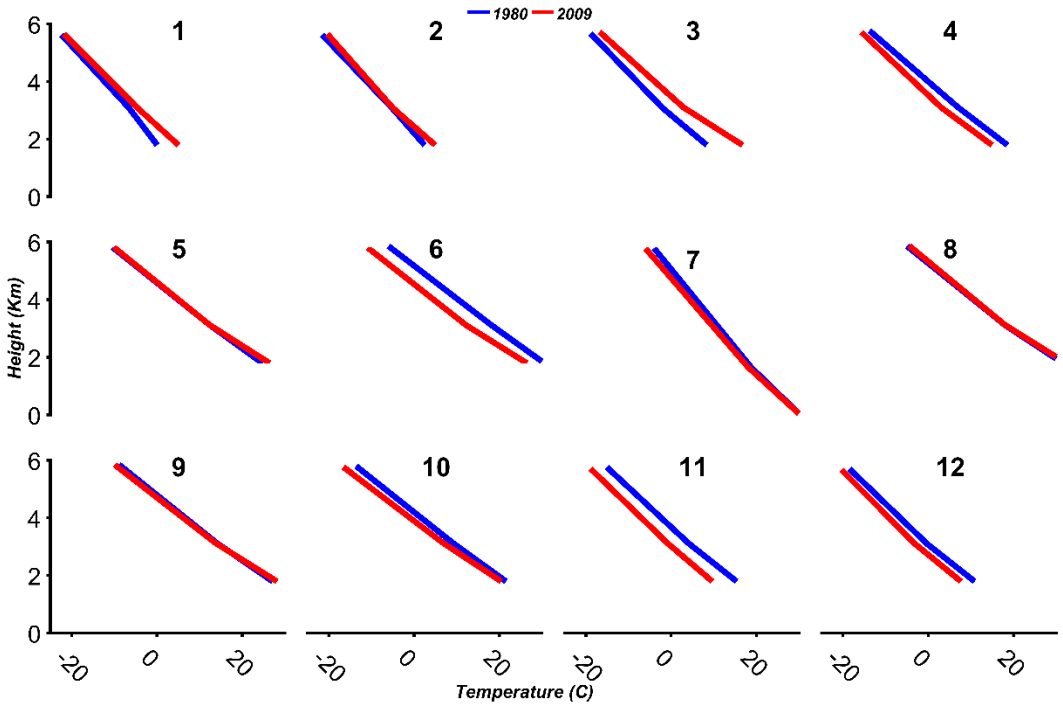


Figure S4. Same as Figure S 4 except for the years 1980 and 2009 at 12-UTC (0430 PM, local time).

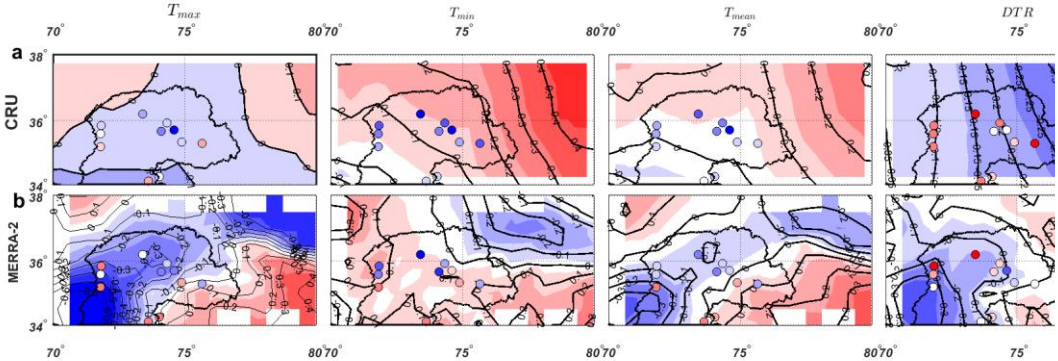


Figure S5. Comparison of temperature trends from CRU and MERRA-2 to the observed datasets. a) Summer seasonal trends of maximum, minimum, and mean temperatures and diurnal temperature range from 1961 to 2011 from station data (circles) and CRU gridded data (contours). b) Same as a) except from 1981 to 2011 using MERRA-2 reanalysis (contours).

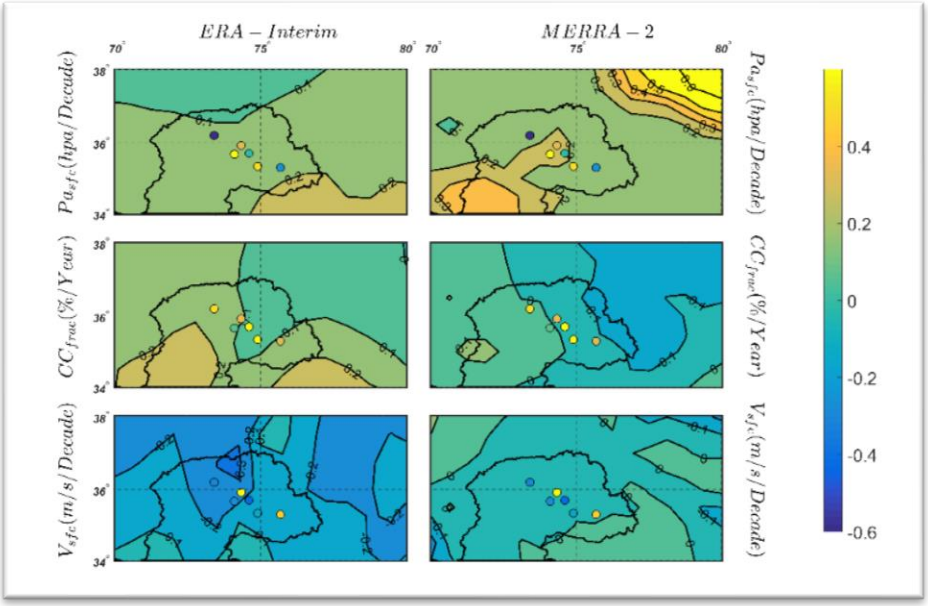


Figure S6. Spatial distribution of ERA-Interim (left panels) and MERRA-2 (right panels) reanalysis as well as in-situ (all panels) observed trends for surface pressure, cloud cover fraction and near surface wind speed.

Table S1. Percentage of missing values to the total time series of maximum and minimum temperature.

	ASTORE	CHILLAS	DIR	CHITRAL
Tx	2.0	-	11.8	8.3
Tn	2.0	0.5	11.8	8.3

Table S2. Same as Table S1, but for other meteorological variables reporting missing values.

	ASTORE	BUNJI	CHILLAS	GUPIS	SKARDU
08-AM DBT	2.0	1.5	-	-	-
05-PM RH%	4.4	-	-	-	-
08-AM, 05-PM Wind Speed	24.5	16.2	26.0	3.4	16.7
08-AM, 05-PM Cloud Cover Fraction, 08-AM, 05-PM Surface Pressure,05-PM DBT, 08-AM RH%	2.0	-	-	-	-

Table S3. Meteorological station and last year of major inhomogeneity found in maximum temperature time series.

	BUNJI	CHILLAS	GILGIT	SKARDU	DIR	GHARI DUPPATA	CHITRAL
Tx	1985	1987	1990	1981	1967	1985	1966

Table S4. Same as Table S3, but for minimum temperature.

ASTORE	BUNJI	CHILLAS	GUPIS	SKARDU	MUZAFARABAD	CHITRAL

Tn	1969	1974	1974	1989	2005	1983	1966

Table S5. Same as Table S3, but for other variables.

	ASTORE	BUNJI	CHILLAS	GILGIT	GUPIS	SKARDU
08-AM DBT	-	1988	1993	1999	1980	2003
05-PM DBT	1992	1986	2002	1985	-	-
08-AM RH%	1966	-	1990	2005	-	-
05-PM RH%	1967	1986	-	-	-	1981
08-AM Wind Speed	1993	1989	1981	-	1981	-
05-PM Wind Speed	-	-		1983	1982	-
08-AM Cloud Cover Fraction	1967	-	1966	-	1977	-
05-PM Cloud Cover Fraction	1966	2005	1966	-	1977	-
08-AM Surface Pressure	2001		1991	1968	2005	1967
05-PM Surface Pressure	2001	-	1991	1966	2005	-
Precipitation	-	-	-	-	1995	-

Table S6. Same as Table S3, but for detrended variables.

	ASTOR E	BUNJI	CHILLA S	GILGIT	GUPIS	SKARD U	DIR	CHITRAL
Tx	-	2005	-	-	-	-	1967	1966
08-AM DBT	-		-	-	2005	-	-	-
05-PM DBT	-	2005	-	-	-	2005	-	-
08-AM RH%	-	-	-	-	-	2005	-	-
08-AM Wind Speed	-	-		-	2005	-	-	-
08-AM Cloud Cover Fraction	2000	-	1966	-	-	-	-	-
05-PM Cloud Cover Fraction	2000	-	1966	-	-	-	-	-
08-AM Surface Pressure	1966	-	-	-	2005	1967	-	-
05-PM Surface Pressure	2001	-	-	1966	2005	-	-	-

Table S7. Seasonal change in Maximum Temperature (Tx) per decade for each site, along with mean regional trend and its standard deviation at seasonal scale. Bold values are statistically significant at 95% confidence interval.

	Tx ($\frac{°C}{10yr}$)				
	WINTER	SPRING	SUMMER	AUTUMN	ANNUAL
ASTOR	0.21	0.37	-0.07	0.26	0.19
BUNJI	0.14	0.11	-0.66	-0.08	-0.12
CHILLAS	0.03	0.14	-0.30	-0.02	-0.03
GILGIT	0.45	0.51	-0.09	0.32	0.30
GUPIS	0.34	0.44	-0.15	0.25	0.22
SKARDU	0.53	0.60	0.14	0.46	0.43
GARHI DUPATTA	0.48	0.63	0.16	0.38	0.41
KOTLI	0.05	0.37	-0.15	-0.12	0.03
MUZAFFARABAD	0.25	0.49	-0.03	0.17	0.22
DROSH	0.22	0.50	-0.03	0.10	0.20
DIR	0.27	0.42	0.07	0.16	0.23
CHITRAL	0.54	0.60	-0.07	0.28	0.32
MEAN ± STD	0.29 ± 0.18	0.43 ± 0.17	-0.1 ± 0.22	0.18 ± 0.18	0.2 ± 0.17

Table S8. Same as Table S7, but for Minimum Temperature (Tn).

	Tn($\frac{℃}{10yr}$)				
	WINTER	SPRING	SUMMER	AUTUMN	ANNUAL
ASTOR	0.15	0.26	-0.15	0.07	0.08
BUNJI	-0.05	-0.03	-0.63	-0.60	-0.33
CHILLAS	0.33	0.18	-0.27	0.05	0.07
GILGIT	0.09	-0.03	-0.40	-0.18	-0.13
GUPIS	-0.26	-0.16	-0.67	-0.44	-0.39
SKARDU	0.13	-0.05	-0.38	-0.40	-0.17
GARHI DUPATTA	0.01	0.11	-0.12	0.04	0.01
KOTLI	-0.05	-0.10	-0.30	-0.26	-0.18
MUZAFFARABAD	0.06	0.13	-0.10	-0.01	0.02
DROSH	0.09	0.20	-0.32	-0.17	-0.06
DIR	0.04	0.00	-0.21	-0.09	-0.07
CHITRAL	0.06	-0.01	-0.37	-0.19	-0.13
MEAN ± STD	0.05 ± 0.14	0.04 ± 0.13	-0.33 ± 0.18	-0.18 ± 0.21	-0.11 ± 0.15

Table S9. Same as Table S7, but for Mean Temperature (Tm).

	Tm($\frac{℃}{10yr}$)				
	WINTER	SPRING	SUMMER	AUTUMN	ANNUAL
ASTOR	0.17	0.31	-0.11	0.16	0.13
BUNJI	0.04	0.04	-0.64	-0.34	-0.23
CHILLAS	0.18	0.16	-0.28	0.02	0.02
GILGIT	0.30	0.24	-0.25	0.07	0.09
GUPIS	0.02	0.14	-0.41	-0.09	-0.08
SKARDU	0.34	0.28	-0.12	0.03	0.13
GARHI DUPATTA	0.24	0.37	0.02	0.21	0.21
KOTLI	-0.02	0.13	-0.23	-0.19	-0.07
MUZAFFARABAD	0.14	0.31	-0.06	0.08	0.12
DROSH	0.14	0.35	-0.17	-0.03	0.07
DIR	0.15	0.21	-0.07	0.04	0.08
CHITRAL	0.27	0.30	-0.23	0.04	0.09
MEAN ± STD	0.16 ± 0.11	0.24 ± 0.1	-0.21 ± 0.18	0 ± 0.15	0.05 ± 0.12

Table S10. Same as Table S7, but for Diurnal Temperature Range (DTR).

	DTR($\frac{°C}{10yr}$)				
	WINTER	SPRING	SUMMER	AUTUMN	ANNUAL
ASTOR	0.06	0.11	0.07	0.19	0.11
BUNJI	0.20	0.14	-0.03	0.51	0.21
CHILLAS	-0.31	-0.03	-0.04	-0.07	-0.11
GILGIT	0.33	0.54	0.31	0.50	0.42
GUPIS	0.61	0.60	0.52	0.70	0.61
SKARDU	0.38	0.65	0.52	0.86	0.61
GARHI DUPATTA	0.46	0.52	0.27	0.34	0.40
KOTLI	0.05	0.47	0.15	0.13	0.21
MUZAFFARABAD	0.17	0.37	0.07	0.17	0.20
DROSH	0.13	0.30	0.29	0.27	0.25
DIR	0.23	0.42	0.28	0.25	0.30
CHITRAL	0.37	0.61	0.32	0.48	0.45
MEAN ± STD	0.22 ± 0.23	0.39 ± 0.22	0.23 ± 0.19	0.36 ± 0.26	0.31 ± 0.21

Table S11. Same as Table S7, but for 08-AM Dry Bulb Temperature.

	08-AM Dry Bulb Temperature ($\frac{°C}{10yr}$)				
	WINTER	SPRING	SUMMER	AUTUMN	ANNUAL
ASTORE	0.18	-0.06	0.44	-0.07	0.06
BUNJI	0.01	0.00	-0.59	-0.54	-0.28
CHILLAS	0.40	0.17	-0.23	-0.14	0.05
GILGIT	-0.05	-0.21	-0.71	-0.54	-0.38
GUPIS	0.13	0.26	-0.23	-0.08	0.01
SKARDU	0.24	0.17	-0.22	-0.27	-0.02
MEAN ± STD	0.15 ± 0.16	0.06 ± 0.18	-0.26 ± 0.4	-0.27 ± 0.22	-0.09 ± 0.19

Table S12. Same as Table S7, but for 05-PM Dry Bulb Temperature.

	05-PM Dry Bulb Temperature ($\frac{°C}{10yr}$)				
	WINTER	SPRING	SUMMER	AUTUMN	ANNUAL
ASTORE	0.33	0.47	-0.01	0.32	0.28
BUNJI	0.14	0.11	-0.66	-0.13	-0.13
CHILLAS	0.17	0.16	-0.21	0.02	0.04
GILGIT	0.42	0.34	-0.25	0.53	0.27
GUPIS	0.23	0.27	-0.35	0.10	0.07
SKARDU	0.34	0.44	0.04	0.11	0.24
Mean ± STD	0.27 ± 0.11	0.3 ± 0.15	-0.24 ± 0.25	0.16 ± 0.23	0.13 ± 0.16

Table S13. Same as Table S7, but for 08-AM Water Vapor Pressure.

	08-AM Water Vapor Pressure $\left(\frac{hPa}{10yr}\right)$				
	WINTER	SPRING	SUMMER	AUTUMN	ANNUAL
ASTORE	0.1	0.1	0.4	0.1	0.2
BUNJI	0.1	0.2	0.6	0.3	0.3
CHILLAS	0.1	0.1	0.4	0.3	0.2
GILGIT	0.1	0.4	0.7	0.3	0.4
GUPIS	0.0	0.2	0.6	0.1	0.2
SKARDU	0.1	0.3	0.6	0.2	0.3
Mean ± STD	0.08 ± 0.04	0.22 ± 0.12	0.55 ± 0.12	0.22 ± 0.1	0.27 ± 0.08

Table S14. Same as Table S7, but for 05-PM Water Vapor Pressure.

	05-PM Water Vapor Pressure $\left(\frac{hPa}{10yr}\right)$				
	WINTER	SPRING	SUMMER	AUTUMN	ANNUAL
ASTORE	0.0	0.1	0.3	0.1	0.1
BUNJI	0.2	0.3	1.1	0.7	0.6
CHILLAS	0.2	0.2	0.6	0.5	0.4
GILGIT	0.2	0.5	1.2	0.5	0.6
GUPIS	0.0	0.3	0.9	0.2	0.3
SKARDU	0.20	0.30	0.80	0.60	0.50
Mean ± STD	0.13 ± 0.1	0.28 ± 0.13	0.82 ± 0.33	0.43 ± 0.23	0.42 ± 0.19

Table S15. Same as Table S7, but for 08-AM Cloud Cover Fraction.

	08-AM CC $\left(\frac{Okta}{10yr}\right)$				
	WINTER	SPRING	SUMMER	AUTUMN	ANNUAL
ASTORE	0.01	-0.05	0.09	0.02	0.02
BUNJI	0.11	0.03	0.18	0.07	0.09
CHILLAS	0.12	0.05	0.21	0.13	0.13
GILGIT	0.03	-0.02	0.14	0.07	0.05
GUPIS	-0.08	-0.15	0.08	-0.02	-0.05
SKARDU	0.04	-0.04	0.15	0.01	0.04
Mean ± STD	0.04 ± 0.07	-0.03 ± 0.07	0.14 ± 0.05	0.05 ± 0.05	0.05 ± 0.06

Table S16. Same as Table S7, but for 05-PM Cloud Cover Fraction.

	05-PM CC$\left(\frac{Okta}{10yr}\right)$				
	WINTER	SPRING	SUMMER	AUTUMN	ANNUAL
ASTORE	0.00	0.00	0.17	0.12	0.07
BUNJI	0.08	0.08	0.28	0.19	0.16
CHILLAS	0.11	0.04	0.18	0.21	0.13
GILGIT	0.01	-0.02	0.20	0.14	0.08
GUPIS	-0.06	-0.16	0.06	0.00	-0.04
SKARDU	0.09	-0.03	0.15	0.11	0.08
Mean ± STD	0.04 ± 0.07	-0.02 ± 0.08	0.17 ± 0.07	0.13 ± 0.07	0.08 ± 0.07

Table S17. Same as Table S7, but for 08 AM Near Surface Wind Speed.

	08 AM Vsfc $\left(\frac{ms^{-1}}{10yr}\right)$				
	WINTER	SPRING	SUMMER	AUTUMN	ANNUAL
ASTORE	-0.03	0.02	0.03	-0.01	0.00
BUNJI	-0.11	-0.11	-0.10	-0.10	-0.10
CHILLAS	0.01	0.00	-0.05	-0.01	-0.01
GILGIT	-0.07	-0.10	-0.06	-0.07	-0.08
GUPIS	-0.17	-0.26	-0.25	-0.17	-0.21
SKARDU	0.00	-0.05	-0.04	-0.04	-0.04
Mean ± STD	-0.06 ± 0.07	-0.08 ± 0.1	-0.08 ± 0.09	-0.07 ± 0.06	-0.07 ± 0.08

Table S18. Same as Table S7, but for 05-PM Near Surface Wind Speed.

	05-PM Vsfc $\left(\frac{ms^{-1}}{10yr}\right)$				
	WINTER	SPRING	SUMMER	AUTUMN	ANNUAL
ASTORE	-0.04	-0.04	-0.16	-0.11	-0.09
BUNJI	-0.11	-0.16	-0.12	-0.10	-0.12
CHILLAS	-0.02	-0.09	-0.14	-0.13	-0.09
GILGIT	-0.08	-0.02	0.01	-0.05	-0.03
GUPIS	-0.15	-0.42	-0.47	-0.32	-0.34
SKARDU	-0.05	-0.15	-0.12	-0.09	-0.10
Mean ± STD	-0.08 ± 0.05	-0.15 ± 0.15	-0.17 ± 0.16	-0.13 ± 0.1	-0.13 ± 0.11

Table S19. Same as Table S7, but for 08-AM Surface Pressure.

	08-AM Psfc ($kPa/10yr$)				
	WINTER	SPRING	SUMMER	AUTUMN	ANNUAL
ASTORE	0.28	0.20	0.34	0.18	0.24
BUNJI	0.23	0.11	0.34	0.17	0.20
CHILLAS	0.46	0.33	0.60	0.45	0.45
GILGIT	0.09	0.06	0.30	0.15	0.14
GUPIS	-0.06	-0.28	-0.17	-0.26	-0.20
SKARDU	0.19	0.22	0.43	0.23	0.25
Mean ± STD	0.2 ± 0.18	0.11 ± 0.21	0.31 ± 0.26	0.15 ± 0.23	0.18 ± 0.21

Table S20. Same as Table S7, but for 05-PM Surface Pressure.

	05-PM Psfc ($kPa/10yr$)				
	WINTER	SPRING	SUMMER	AUTUMN	ANNUAL
ASTORE	0.40	0.21	0.45	0.27	0.33
BUNJI	0.33	0.19	0.49	0.25	0.30
CHILLAS	0.42	0.31	0.68	0.34	0.43
GILGIT	0.16	0.07	0.43	0.22	0.21
GUPIS	-0.03	-0.28	-0.05	-0.20	-0.15
SKARDU	0.16	0.13	0.27	0.09	0.15
Mean ± STD	0.24 ± 0.17	0.11 ± 0.21	0.38 ± 0.25	0.16 ± 0.2	0.21 ± 0.2

Table S21. Same as Table S7, but for Net Radiations.

	Rn $\left(\frac{MJm^2D^{-1}}{10yr}\right)$				
	WINTER	SPRING	SUMMER	AUTUMN	ANNUAL
ASTORE	0.01	0.00	-0.11	-0.01	-0.01
BUNJI	0.00	-0.03	-0.20	-0.03	-0.05
CHILLAS	0.00	-0.01	-0.16	-0.03	-0.05
GILGIT	0.00	0.03	-0.12	-0.01	-0.02
GUPIS	-0.01	0.10	-0.02	0.00	0.02
SKARDU	0.01	0.04	-0.09	0.01	-0.01
Mean ± STD	0 ± 0.01	0.02 ± 0.05	-0.12 ± 0.06	-0.01 ± 0.02	-0.02 ± 0.03

Table S22. Same as Table S7, but for Potential Evapotranspiration.

	ETP $\left(\frac{mmD^{-1}}{10yr}\right)$				
	WINTER	SPRING	SUMMER	AUTUMN	ANNUAL
ASTORE	0.00	0.01	-0.12	-0.04	-0.03
BUNJI	-0.06	-0.13	-0.29	-0.16	-0.15
CHILLAS	0.01	-0.05	-0.24	-0.12	-0.10
GILGIT	-0.03	-0.06	-0.16	-0.06	-0.08
GUPIS	-0.06	-0.20	-0.47	-0.22	-0.23
SKARDU	0.00	-0.04	-0.16	-0.08	-0.07
Mean ± STD	-0.02 ± 0.03	-0.08 ± 0.07	-0.24 ± 0.13	-0.11 ± 0.07	-0.11 ± 0.07

Table S23. Same as Table S7, but for Precipitation.

	Precipitation $\left(\frac{mmD^{-1}}{10yr}\right)$				
	WINTER	SPRING	SUMMER	AUTUMN	ANNUAL
ASTORE	0.12	-0.14	0.08	-0.01	0.01
BUNJI	0.06	-0.01	0.08	0.01	0.03
CHILLAS	0.06	0.02	0.08	0.00	0.04
GILGIT	0.03	0.02	0.04	0.01	0.02
GUPIS	0.08	0.18	0.12	0.05	0.11
SKARDU	0.13	0.06	0.04	0.00	0.06
Mean ± STD	0.08 ± 0.04	0.02 ± 0.1	0.07 ± 0.03	0.01 ± 0.02	0.05 ± 0.04

APPENDIX B: DEVELOPMENT OF HIGH RESOLUTION DAILY PRECIPITATION AND TEMPERATURE DATA PRODUCT FOR PAKISTAN

(Submitted to International Journal of Climatology)

Furrukh Bashir

Department of Hydrology and Atmospheric Sciences, University of Arizona, Tucson, AZ, USA

Pakistan Meteorological Department, Islamabad, Pakistan

Xubin Zeng, Hoshin Gupta, Pieter Hazenberg

Department of Hydrology and Atmospheric Sciences, University of Arizona, Tucson, AZ, USA

ABSTRACT

The quality of observation-based gridded hydro-meteorological data sets depends on the quantity and quality of the input observations, and existing gridded data sets for Pakistan are not representative of the actual density of available observations. Here, we report on the development of a gridded daily 0.05º×0.05º maximum and minimum temperature and precipitation data set for Pakistan, named PAK-HYM-1.0, based on data recorded at 67 sites over the period 1960-2013. We give special attention to the methods for handling missing data, and for point to grid interpolation, and examine the sensitivity of the interpolation algorithm to number of observations provided. Geo-statistical analysis is used to infer the desirable spatial resolution for a gauging-network to obtain estimates of precipitation in different areas of Pakistan. Compared to other gridded products, PAK-HYM-1.0 has finer spatial and temporal resolutions, and extends over a longer period. More importantly, it is based on data from two to four times more meteorological stations than are used in the development of other products. Among other benefits, the data set is useful for operational monitoring of climatic events (such as wet and dry periods) at daily to monthly scales. The product will be updated periodically using additional observations and new hydro-meteorological variables over the coming years.

Keywords: Gridded Observations, Precipitation, Temperature, Hydro-meteorological dataset, Pakistan.

1. INTRODUCTION

Environmental impact models are now commonly used to assess the potential consequences of climate change (e.g. impacts of warming on crop yield or land cover through change in rainfall). Such models require spatially and temporally continuous data. For example, to analyze the impacts of large scale circulation or teleconnections of a region it is common to use area-average climatic indices (Chang *et al.*, 2004; Lau and Wu, 2001; Mariotti, 2007) estimated from spatially continuous gridded data sets that are developed using objective analyses of the observations available at irregularly-spaced points (Chen *et al.*, 2002, 2008). Moreover, since

climate models generally simulate representative areal averages of relevant processes at regional scales (Osborn and Hulme, 1998), evaluation of their simulations necessitates a direct comparison between model simulated fields and gridded observations at the same spatial scale. Operational applications such as climate disaster monitoring by different National Hydro-Meteorological Services (NHMS) also use gridded temperature and precipitation data as an input. At present, the major data sources for gridded products are in-situ observations, satellite retrieved information and outputs from numerical models, each of which has its strengths and weaknesses (Xie and Arkin, 1997).

The most rudimentary, common, direct, and longest-lived precipitation quantification method for precipitation is the in-situ use of rain gauges. At the gauge scale, the "catch" of a gauge is affected by wind flow around the orifice (Duchon and Essenberg, 2001), resulting in under-catch at low rainfall intensities and higher wind speeds (Nešpor *et al.*, 2000). Despite such errors, however, networks of rain gauges arguably remain the most accurate way to measure rainfall at the land surface, and conventional observations through surface gauge networks provide the most direct measurements of precipitation (Kidd and Huffman, 2011). Estimates of air temperature provided by station networks have similar causes of uncertainty; i.e., they may be affected by insufficient station densities to resolve the spatial variability of temperature, particularly in complex terrain [*Stahl et al.*, 2006].

Gridded data products are often based on historical observational records, mainly collected at the land surface (by irregularly spaced meteorological observatories) and interpolated to a gridded format that provides "best" estimates of relevant climatic variables at locations away from the observation collection sites. Therefore, such products are useful for estimating the values of variables under consideration in data-sparse locations (Haylock *et al.*, 2008). Since the original data used to develop different gridded data products is often the same, although products may differ based on how the station records are filtered, interpolated, and homogenized the general climatological agreement of different gridded data products is not surprising (Tapiador *et al.*, 2012). Note that most of the gridded precipitation and temperature products consider topographic effects, and this may introduce large errors due to disparities in the elevation model from the original

hypsometric heights. Nevertheless, in areas where observation densities are high enough to overcome the discrepancies of the elevation model, such topographic errors may be masked.

A second major source of temperature and precipitation estimates at near global scales is satellite-based observation, now approaching nearly four-decades of availability. While satellite based products are superior to their ground observed counterparts in regard to spatial coverage, they are subject to retrieval errors and biases. Satellite-based estimates of precipitation draw upon information provided by visible, infrared, and passive and active microwave systems, using a variety of techniques that differ in performance at different regional and temporal scales (Kidd and Huffman, 2011). In brief, the quality of satellite-based retrievals of precipitation estimates are limited by the following factors (Barrett, 2000):

- The temporal, spatial and spectral characteristics of the satellite observing systems.
- The varied, and variable, interactions between electromagnetic radiation and the environment particularly over land areas.
- Operational techniques that employ fixed rain/no rain IR temperature thresholds and constant rain rates for all parts of the globe (Arkin *et al.*, 2017).

Similarly, air temperature estimates from satellites are limited by intrinsic characteristics of the satellites and availability of limited in-situ meteorological data.

A third major approach used to develop gridded temperature and precipitation products is the use of Numerical Weather Prediction (NWP) models and Retrospective-analysis (reanalysis) to estimate precipitation in forecast mode based on temperature and precipitation information derived from assimilated observations. The efficiency of precipitation estimation via reanalysis is typically evaluated by comparison with gridded observed products, and any tendency to under/overestimate precipitation is corrected to match the gridded observations (Dee *et al.*, 2011). Hence, the gridded observations are usually considered as the first approximation of the truth.

Each product, whether gridded observed, satellite retrieved or forecasted by numerical models, has its own limitations, as described in Supporting Information (S.I) and Table S1. In general, gridded observed data sets are considered to be superior if the gauge density is higher. However, the precipitation gauge network of Pakistan is not well represented in available gridded products for the region. Accordingly, we present here a daily, in-situ observation based, 0.05º×0.05º gridded temperature and precipitation data set for Pakistan, for the period of 1960-2013. We name it PAK-HYM-1.0, that is an abbreviation of Pakistan and Hydrometeorology, and 1.0 indicates that it is the first version. This data set is developed by utilizing data from 67 meteorological observatories across Pakistan. This number of observation sites incorporated is 2 to 4 times higher than that used in other similar products available for this region.

The study area is described in section 2. The data used in this study and the methodology for development of the gridded-products are described in section 3. Section 4 presents the results, and our conclusions are discussed in section 5.

2. STUDY AREA AND CLIMATOLOGY

2.1 STUDY AREA

Pakistan lies in the southwestern portion of South Asia, between 23.45ºN to 36.75ºN latitude and 61ºE to 75.5ºE longitude. The country is bordered by the Hindukush, Karakoram and Himalayan (HKH) mountain ranges to the north, and shares common borders with China (595 km) to the northeast, India (909 km) to the east, and Afghanistan (2240 km) to the northwest. The Arabian sea lies to the south and Iran to the west. Nearly 55% of Pakistan is covered by mountains and plateaus, and the remaining 45% is plains and deserts (Figure B. 1).

The boundaries of Pakistan are oriented in a southwest to northeast direction. The northern and western parts are hilly areas, while the central and southeastern parts are plains or deserts. Topography, see digital elevation model (DEM) presented in Figure B. 1, has a profound impact on the distributions of temperature and precipitation across the country. Elevations vary from sea level in the south to above 8000 meters above sea level (masl) in the Karakoram mountains to the

North. 365-day climatology of Maximum Temperature, Minimum Temperature and Precipitation of several sites in different parts of Pakistan are presented in Figure B. 1.

2.2 CLIMATOLOGY OF PAKISTAN

Starting from the northwestern region, we examine the climatology of different locations as we rotate counterclockwise through the country. The Drosh (1) and Chitral (2) sites are located in the hilly northwest, where interception of westerly disturbances by the high-altitude mountains causes precipitation mainly in the winter and spring, with summer and autumn being relatively dry seasons. The mean maximum temperature rises to 40ºC in dry conditions, exhibiting less variability and expansion of the Diurnal Temperature Range (DTR). Monsoon penetration to this part of the country is very limited.

The next two sites, Cherat (3), and Parachinar (4) are in the northwestern part of the country at a lower elevation than locations (1) and (2). In addition to wet winter and spring seasons, these sites receive a substantial amount of monsoon rainfall in summer. The greater availability of moisture reduces daytime temperatures, but the DTR is high in dry periods. Further south in a counterclockwise direction, Quetta (5) is situated a bit higher then Sibbi (6), and is accordingly colder. Quetta receives more winter rainfall than Sibbi, but less in the summer monsoon season. Sibbi is surrounded by mountains on all sides except the southeast, making it a rain shadow location with dry and extremely warm conditions prevailing throughout most of the year.

Further south, Nokkundi (7) is selected as a westernmost location of Pakistan. It receives little annual precipitation, occurring mainly in winter, and is mainly hot and dry. Jiwani (8) is a coastal location strongly influenced by proximity of the Gulf of Oman, and the Arabian Sea. Here, temperatures are relatively consistent, with smaller year around DTRs than any other location in Pakistan. It receives more annual precipitation then Nokkundi (located to the north), mostly during the winter season. Although, Jiwani receives little summer monsoon precipitation, the daytime temperatures are lower mainly due to cloudiness.

Karachi (9), one of the southernmost locations, is a coastal city that receives substantially more summer monsoon rainfall than its western neighbor (Jiwani). Summer temperatures decline

with onset of the monsoon, and DTR decreases. In winter a continental effect dominates with little rainfall. The southeastern most locations of Badin (10), Chhor (11) and Kanpur (12) receive the dominant share of their annual precipitation in monsoon season, which also results in a drop in temperature.

Moving north to Jhelum (13) and Lahore (14), copious rainfall is experienced during the monsoon season together with increased precipitation in other seasons also. The same trend continues for Islamabad (15). In this area, the autumn and spring seasons are relatively hot. Day temperatures drop during the monsoon and the DTR decreases. Muzaffarabad (16) is a mountainous location where the monsoon is largely obstructed, however, it receives more winter season rainfall than Drosh (1) and Chitral (2). Going farther north, Gupis (17) and Skardu (18) receive the main contribution of their annual rainfall in winter, along with sporadic episodes of summer rainfall.

Based on this analysis, we can divide Pakistan into three hydrometeorological regions: 1) the north, northwestern and western mountainous areas that receive mainly winter precipitation, 2) the south and southeastern areas that receive mainly summer monsoon precipitation, and 3) the remaining areas that receive varying amounts of precipitation in both summer and winter, such as the River Indus plains in central Pakistan. The spring and autumn seasons are generally dry throughout the country. Spring is warm in the plains and summer is warm in the mountainous areas. Dry locations away from the Arabian Sea are the warmest.

3. GENERATION OF A GRIDDED DATA PRODUCT FOR PAKISTAN

3.1 PRECIPITATION AND TEMPERATURE DATA

Daily precipitation and maximum and minimum temperatures are available for 67 sites located throughout Pakistan. A list of the sites, along with their location, elevation and length of data availability is presented in Table S2. All of these stations are operated and quality-controlled by the National Hydro-Meteorological Service (NHMS) of Pakistan (PMD); data can be requested at http://cdpc.pmd.gov.pk/home.htm.

Upon examination of the data, we found incidences when the maximum temperature is recorded as being less than the minimum temperature. We identified all such events and dropped

those values that did not fall within four standard deviations of the climatological monthly mean (via an iterative process). Further, in some instances, temperatures were found to be reported in degrees Fahrenheit (ºF) instead of degrees Celsius (ºC), or precipitation values were reported in inches instead of millimeters. We corrected such inconsistencies by painstaking subjective inspection of the time series of each location. Figure B. S1a presents the total number of observations (excluding missing values) available for precipitation, maximum and minimum temperature over the period 1960-2013. Note that the total number of observations has been increasing with time, and the largest number of available observations is for the last decade of the time series. Figure B. S1b presents the percentage of missing values for each site. Most sites have ~5-10% missing values, and only a few sites have greater than 10% missing data.

3.2 ACCOUNTING FOR MISSING DATA

To address the problem of missing values (imputation), we used the available data to estimate the annual climatology for all of the variables at each site for each of the 365 days of the year. We further smoothed the estimated climatologies by applying a cubic smoothing spline of 0.1% variability, where the 365th value is placed before the 1st value to obtain a full 365-day smoothed climatology. In a few instances the raw 365-day climatology of precipitation (at a very dry location) had a small value for several days and cubic spline smoothing resulted in some negative values; such values were replaced by 0.01 to minimize contamination of the site's climatology by unrealistic values. As an alternative to cubic spline smoothing, we also tested the use of Fourier truncation with accumulation of its six leading modes to remove the high frequency signal from the raw climatology as suggested by Xie *et al.* (2007); however, this resulted in overestimation of values at the beginning and the end of the curves, and so the results are not shown here.

To fill in reasonable estimates of values for the missing data (imputation), we developed a modified version of the inverse distance weighting method (IDWM; e.g., Teegavarapu and Chandramouli, 2005), referred to here as the inverse climatology deviation weighting (ICDWM) method (see Equation 1). In this approach, we estimate the climatological deviation from the target

site of all neighboring sites for each Julian day for which data values are missing. Using inverse weighting we assign higher weights to sites that have a small deviation of daily climatology for the specific day under consideration, and low weights to the ones having large deviation. Instead of using absolute values, we use the ratio (for precipitation) or difference (for maximum and minimum temperature) of the target date's value to the respective day's climatology of the neighboring sites to estimate the error term that will be multiplied/added to the target site's smooth climatology (for the respective Julian day) to replace the missing value of that date, as explained in equation (1).

$$\epsilon(x) = \left(\sum \epsilon_j / w_j\right) / \left(\sum 1/w_j\right) \quad (1)$$

where $j = 1,2,3, \ldots, n$ presents neighboring sites, ϵ is the error term or the deviation (or scaling factor) of the variable at a specific site for a given date from its long-term climatology (μ) as presented in equations (2) and (3).

$$y_{(j,t)} = \mu_{(j,t)} + \epsilon \quad \text{or} \quad y_{(j,t)} = \mu_{(j,t)} \times \epsilon \quad (2)$$

$$\epsilon = \frac{\sum \frac{y_{(j,t)} - \mu_{(j,t)}}{w_{(j,t)}}}{\sum \frac{1}{w_{(j,t)}}} \quad \text{or} \quad \epsilon = \frac{\frac{\sum y_{(j,t)}}{\mu_{(j,t)} \times w_{(j,t)}}}{\sum \frac{1}{w_{(j,t)}}} \quad (3)$$

3.3 CROSS-VALIDATION OF ICDWM IMPUTATION

The performance of this algorithm was tested by use of leave-one-out cross-validation, by removing a whole year at a time from the time series for each location, and replacing it with estimates generated via ICDWM using observations available from neighboring locations. Results are depicted in Figure B. 2 via a Taylor diagram (Taylor, 2001), which provides a concise statistical summary of how well patterns match with each other in terms of correlation, root-mean-squared (RMS) difference and the ratio of their variances, as expressed via the following formula:

$$E'^2 = \sigma_e^2 + \sigma_r^2 - 2\sigma_e \sigma_r R \quad (4)$$

where R is the correlation coefficient between the estimated and reference time series, E' is the centered RMS difference between the fields, and σ_e^2 and σ_r^2 are the variances of the estimated and reference time series, respectively.

3.4 DEVELOPMENT OF THE GRIDDED CLIMATOLOGY AND DAILY GRIDDED DATA PRODUCT

A 365-day gridded climatology product for each grid-point lying within the political boundary of Pakistan was developed by interpolation of the in-situ observation data using a two-step methodology. First, a local climate-elevation relationship was developed for each target grid-point to model the strong dependency of precipitation and temperature distributions on elevation (Daly *et al.*, 2008), based on the climatological value of precipitation or temperature; this was done for each day via linear regression against elevation (characteristics of DEM are explained in S.I). Products were developed for precipitation, maximum, and minimum temperature on a 0.05º latitude-longitude grid lattice.

However, in cases where the target grid-point is located far high or low in relation to participating in situ location, this approach will not provide a suitable estimate and, instead, the simple mean of the neighboring sites was adopted as a first estimate of the truth. In the second step, the unexplained variation of the fitted model at the target grid point was estimated using the residual errors (minimum mean squared residuals) of the neighboring sites using the Spheremap algorithm (Willmott *et al.*, 1985). Spheremap uses an empirical angular distance weighting scheme (Shepard, 1968) on a 0.05° regular latitude-longitude grid, based on the simple assumption that the value (noise) of each nearby point influences the estimate of the associated grid-point value (noise) by an amount inversely proportional to the distance between them. The influence is described by weights such that the sum of the predictor weights is 1; details of the interpolation scheme are provided in Willmott *et al*, (1985), and S.I.

After computing the gridded climatology for each day (of 365-days), the 365-day climatology at each grid point was again smoothed using cubic splines, as described earlier.

Next, for each calendar date, the ratio/difference of the daily observations to their 365-days climatology was interpolated (Yatagai *et al.*, 2012) using the Spheremap interpolation method (Willmott *et al.*, 1985) and then multiplied/added to the gridded (365-days) climatology at each grid-point for the respective Julian day.

3.5 SENSITIVITY OF THE INTERPOLATION ALGORITHM TO OBSERVATION DENSITY

The sensitivity of the interpolation algorithm to the number of observations utilized (observation density) was tested as follows. To test the robustness of the interpolation scheme, a subset of the data (up to 70% of the sites) was held back for use as evaluation sets, and the interpolation algorithm was applied to the remaining data. Averaging the squared difference yields an overall measure of estimation accuracy. To improve robustness of the approach, we iterated 1000 times using a randomization scheme in which a random number of data sites were selected with replacement.

We compared the mean annual precipitation of each year from PAK-HYM-1.0 to the mean annual precipitation developed through randomized subsets and computed the RMSE for each year. The average RMSE over the entire duration (1960-2013) was adopted as an indicator of robustness, and the procedure was repeated 1000 times with random combinations of sites (20-67); results are presented in Figure B. 6a. Similar analyses for annual mean maximum and minimum temperature are presented in Figures B. 6b and B. 6c.

To describe the statistical variation and structure of the in-situ rain gauge observations, we use an experimental 'Variogram' (Berne *et al.*, 2004), which characterizes the differences of variable magnitude between spatial observations as a function of their respective distance. Assuming the spatial field of the variable under consideration is a realization of a random function y, the structure of the function is characterized by the variogram γ:

$$\gamma(\mathrm{d}) = \frac{1}{2n(d)} \sum_{d_{ij}=d} \left(y_i - y_j\right)^2 \qquad (5)$$

where $\gamma(\mathrm{d})$ is the semi-variance defined over observations y_i and y_j lagged successively by distance d. The variogram provides a measure of dissimilarity of observations at a distance d, whereas the covariance is a measure of similarity at same distance, estimated as follows:

$$C(\mathrm{d}) = \frac{1}{N}\sum_{ij=1}^{N}\left(y_i - \bar{y}_i\right)\left(y_j - \bar{y}_j\right) \qquad (6)$$

Three aspects of a variogram are the "nugget" effect, the "sill" and the "range". The nugget is a possible discontinuity at the origin, arising due to coarseness of observation resolution or measurement errors. The sill occurs when a random function has a defined variance at a given range. In this case, the sill corresponds to the variance (standardized to 1 for the climatological variogram) and the range to the decorrelation distance. Estimates of the nugget, sill and range were used to characterize the spatial structure of precipitation over Pakistan. We estimated the "climatological variogram" and "climatological covariance" for mean summer and winter precipitation (1960-2013) to describe the seasonal variability of the spatial characteristics of precipitation over Pakistan and the effectiveness of rain gauges to capture them (Figure B. 7). Based on this, we suggest a suitable number for the rain gauge density required to adequately resolve the spatial variability in different parts of Pakistan (Table 1).

4. RESULTS AND DISCUSSION

4.1 MISSING DATA IMPUTATION

As discussed in section 3.2, missing data in the observational record was addressed using the ICDWM method. To investigate robustness of the ICDWM approach, we applied the leave-one-out cross-verification method and compared the reconstructed time series with the original one. For a given site, all of the observations for year 2011 were replaced by values reconstructed using the time series data from neighboring sites. In addition, the monthly means of maximum and minimum temperatures and monthly accumulated precipitation were computed for both the observed and reconstructed time series at each site. A statistical summary of spatial pattern matching is presented in Figure B. 2 as a Taylor Diagram.

The results show that monthly mean maximum temperatures are reproduced very well, with correlation coefficient centered around 0.99 and centered root mean squared error (CRMSE) around 0.1 to 0.2. The reconstructed data explains nearly all the variance present in the original data. Meanwhile, the correlation coefficient of reconstructed to observed minimum temperatures is only 0.05 lower than that obtained for maximum temperature reconstruction. This slight decrease is explained by the larger number of missing values in the original data from which the smoothed climatology was estimated (see Figure B. S1). The correlation coefficient obtained for monthly accumulated precipitation varies throughout the year, ranging from 0.3 in May to 0.92 in April of the selected year (i.e., 2011). Overall, the reconstructed monthly accumulated precipitation time series show less variability than the original data, due mainly to the fact that the interpolation algorithm overestimates small values (it sometimes generates small magnitude events when there is no event in original data) and underestimates extreme events. Nevertheless, keeping in mind the stochastic nature of the precipitation data, the performance of the algorithm in reconstructing precipitation is reasonable.

To evaluate the quality of reconstruction of the gridded precipitation time series, the precipitation events were divided into the following three categories:

1) Cat-I: Low Intensity Events ($\leq 70^{th}$ percentile)
2) Cat-II: Moderate Intensity Events ($> 70^{th}$ percentile & $\leq 90^{th}$ percentile)
3) Cat-III: Extreme Intensity Events ($> 90^{th}$ percentile)

These categories were further subcategories into Type-I, Type-II and Type-III events, as follows:

a) Type-I: False-positive or spurious precipitation events that are not present in observations but generated by algorithm due to neighborhood effect.
b) Type-II: False-negative events that are present in observations but couldn’t reproduced by the algorithm.

c) Type-III: Events that exists in observation and are effectively estimated.

Among the low intensity events (Cat-I) the probability of Type-I events (spurious precipitation) was found to be the highest (Figure B. 3a). The spatial map of probability of such events indicates that they are more likely to happen in areas having higher observation network densities, where meso- and micro-scale events occurring in neighboring locations may result in the generation of pseudo-events for a given location. This causes a high number of such events to be reported in the low intensity portions of the precipitation data, and this happens all over the country regardless of seasonal dependency (Figure B. 3a1).

The mean probability of Type-II events in Cat-I is small, and its spatial distribution indicates that such events (when observed precipitation is not effectively reproduced) occur mainly in areas with low observation densities (southwestern parts of Pakistan) where Cat-I events are so localized that neighboring sites do not report any information about such events, whereas, areas with high observation density have very low probability for Type-II events. This suggests that higher densities of the observational network may help to improve the estimates in such regions (Figure B. 3a2). Finally, the mean probability of Type-III in Cat-I is low, mainly because the probability space is largely occupied by the former two subcategories (Figure B. 3a3).

For the moderate intensity events (Cat-II, Figure B. 3b), the mean probability of Type-I events is lower than that of Cat-I and, in contrast, such events mainly occur in the areas having relatively sparse observation networks (central Pakistan and southwestern areas Figure B. 3b1). Such pseudo precipitation events occur due to precipitation being observed at the “neighboring” locations. The mean probability of Type-II error increases in Cat-II; however, such events are localized to the monsoon dominated southeastern locations and, as such, are less affected by micro- and meso-scale precipitation events observed at neighboring locations (Figure B. 3b2). The mean probability of Type-III events is also larger than for Cat-I (CAT-II, Figure B. 3b3).

Finally, the mean probability of Type-I error in extreme intensity events (CAT-III) is smaller and limited mainly to very sparsely observed areas (Figure B. 3c, c1). The mean probability

of Type-II events is very small (Figure B. 3c2), whereas that of Type-III events is further improved (Figure B. 3c3).

4.2 COMPARISON OF PAK-HYM-1.0 WITH OTHER PRODUCTS

The newly created precipitation product was compared with other available (observed and satellite based) precipitation products (listed in Table S1) for the summer period, when spatial variability is highest (Figure B. 4a-h) due to the monsoon. By using a common 0.5° grid lattice for the evaluation, we can compare the magnitudes of mean summer precipitation across all of the products.

The mean summer precipitation fields from PAK-HYM 1.0, and other leading products (Table 1) are presented in Figure B. 4a-i. Though not all products cover the exact same period of time, their inter-annual seasonal averages are still comparable. Summer season precipitation is dominated by monsoon influences, and therefore may exhibit more spatial variability than in other season due to its mesoscale convective nature.

The results indicate that the CRU product (Figure B. 4e) is most similar to PAK-HYM-1.0, with a correlation value of r^2 = 0.92, low CRMSE, and relatively better explanation of the variance. However, CRU seems to overestimate precipitation in the mountainous areas as compared to the in-situ observations. In terms of correlation, the other products that also appear to perform well are TRMM-3B43, GPCC and Aphrodite. Among those, GPCC has a better explanation of the variance (with smallest CRMSE). Overall, PAK-HYM-1.0 can be considered superior to these products, because of its higher gauge densities, spatial resolution (0.05º vs 0.5º), and temporal resolution (daily vs monthly).

The APHRODITE product is of particular interest because it, like PAK-HYM-1.0, is a daily product. We therefore compared its performance to that of PAK-HYM-1.0 on seasonal and daily scales (see Figure B. 5). Note that Pakistan suffered from severe flooding in 1992, and so the month of June 1992 was therefore selected for the comparison. It is clear from Figure B. 5a that APHRODITE was unable to indicate rainfall in areas where it lacks access to observations, such as central Pakistan, it overestimated rainfall (as compared to observations) in areas of Khyber

Pakhtunkhwa (KPK), and underestimated rainfall in the southeastern areas of Sindh. At the daily scale, on June 24, 1992, it is completely misses a significant event in central Pakistan (southern Punjab) and underestimates rainfall in the southeastern areas (Sindh). Moreover, PAK-HYM-1.0 provides better temporal coverage of precipitation in Pakistan (up to 2013 as compared to 2007 of APHRODITE).

Comparing CHIRPS (monthly, 0.5º) with PAK-HYM-1.0 we see good correlations and variance explained, along with small CRMSE, at the seasonal time scale. CHIRPS reports higher precipitation amounts in areas away from the in-situ observation locations, and especially in the northern mountain ranges. In the central and southern regions, it sometimes does not indicate precipitation events; e.g., for June 2010 (Figure B. 5c), a major flood year in Pakistan, CHIRPS does not indicate any rainfall during the entire month, which is a considerable discrepancy. Note also that CHIRPS covers only the period 1981- present and therefore does not provide a long term history of precipitation for the country.

The CPC (monthly, 0.5º) product ranks next in terms of correlation, but explains less of the variability and has poorer CRMSE performance. It mainly reports lower rainfall values in regions away from the in-situ observations. Overall, its spatio-temporal resolution is poor. The CMORPH (daily, 0.25º) product ranks next in terms of correlation, but has a relatively large CRMSE and tends to “overestimate” in the northern and western mountainous ranges. However, as shown in Figure B. 5g for August 03, 2010, CMORPH sometimes indicates no rainfall at locations where intense rainfall events are indicated by the ground based observations; this may be due to those rainfall events being intense and short lived, occurring in the temporal gaps between satellite sample observation times. Finally, the PERSIANN-CDR product has the lowest correlation ranking, lower variability explained and highest CRMSE. Overall, it generally seems to “underestimate” in areas of high precipitation and “overestimate” in areas of low precipitation.

Our comparison of PAK-HYM-1.0 with other products, on the basis of both individual events and at seasonal time-scales, highlights the importance of having higher observation densities for the development of gridded products of this kind. Collectively, the observations-based products tend to perform better than the satellite-based products. The new PAK-HYM-1.0 product benefits

from input information (in-situ observations) that is of better spatial and temporal resolution, and spanning a longer period of time.

4.3 PERFORMANCE OF THE INTERPOLATION ALGORITHM

The sensitivity of the interpolation scheme to availability of observations was estimated by random subsampling combined with bootstrapping as discussed in Section 3.5.

Figure B. 6 presents the results of the analysis. When only 20 stations are used, the average value for RMSE (Root Mean Squared Error) of the estimated mean annual precipitation (compared to when all observations are used) is around 0.6 mm/day (Figure B. 6a); as the number of stations is increased, the RMSE declines towards zero, with a more rapid rate of decline above ~53 stations. Similar results can be seen for maximum daily temperature (Figure B. 6b) and minimum temperature (Figure B. 6c). These results indicate that the interpolation algorithm is performing in a consistent manner.

To improve our understanding of the quality of the gridded precipitation data product presented here, we examine the summer/winter climatological semi-variograms (Berne *et al.*, 2004) of available in-situ precipitation observations. This helps to provide a quantitative assessment of desirable spatial resolution for in-situ rainfall gauging. North-South (N-S) and East-West (E-W) directional semi-variograms are presented in Figure B. 7 separately for the summer and winter seasons. Notice that all of the directional semi-variograms indicate relatively smooth decays in spatial-correlation, with most of that decay occurring at distances between 100 km and 190 km, the latter being the range (de-correlation distance) of the semi-variogram, estimated as the distance from the origin to the 95% of the sill.

These semi-variograms (Figure B. 7) characterize the dominant structure of directional variability in the summer and winter seasons (the dominant rainy seasons) of Pakistan. Figure B. 7a1 indicates a relatively smooth increase in variance (dissimilarity) in the North-South direction, associated with the pattern of summer monsoon rainfall. During this season, winds from the Arabian Sea push moisture northward, giving rise to the smooth correlation structure (see precipitation climatology in Figure B. 1). On the other hand, Figure B. 7a2 indicates some E-W

structures at scales of 50-100 km, associated with dissimilarities between sites at those distances apart; see Figure B. 4i, which indicates an E-W gradient of precipitation between 32°N to 35°N that can be considered the main cause of these dissimilarities.

Similarly, winter season precipitation in Pakistan is dominated by the extratropical storms, also known in India and Pakistan as the "western disturbances" (see precipitation climatology in Figure B. 1), giving rise to the relatively smooth E-W semi-variogram (Figure B. 7b2). Because western disturbances typically do not penetrate the southern regions, the N-S directional semi-variogram indicates large dissimilarities between different sites (via the nugget effect).

Berne *et al.* (2004) have presented an approach which can be used to estimate the representative (circular) region that can be associated with any given rain gauge based on knowledge regarding the range of the semi-variogram. The estimate uses the equation:

$$S_r = \pi \left[\frac{\Delta}{2}\right]^2, \text{ with } \Delta \in \left[\frac{1}{4}r, \frac{1}{2}r\right], \qquad (7)$$

where S_r is the effective area of the representative circular region, having diameter Δ, which indicates the spatial resolution (Berne *et al.* 2004) and is related to the range r (de-correlation distance) of the semi-variogram. Consistent with the fact that Pakistan can be treated as three different hydrometeorological regions (see Section 2), we estimated the range of spatial correlation in each of the three parts of the country based on the following classification (See Table 1 for details);

i) Northern Pakistan (Stations > 35°N)

ii) Middle Pakistan (Stations between 30°N and 35°N)

iii) Southern Pakistan (Stations < 30°N)

In each of these regions, we estimate the de-correlation distance r via a randomization process in which a specified number of stations are selected randomly with replacement (see Figure B. 7) that have mean inter-gauge distance nearly equal to the mean inter-gauge distance of all sites in each of the above-mentioned part of the country. Results are computed separately for summer

(Figures B. 7a3-a5) and winter (Figures B. 7b3-b5). In these figures, the bar and whisker plots indicate the average decorrelation length obtained for a given number of selected stations. The results indicate that r is approximately 75 km in the North, 210 km in the Middle, and 275 km in Southern Pakistan.

To utilize Equation 7, we must select a value for the spatial resolution Δ. As per Berne *et al.* (2004), the ratio $\frac{\Delta}{r}$ should be on the order of $\frac{\Delta}{r} \in \left[\frac{1}{4}, \frac{1}{2}\right]$. Plugging these values into Equation (7), we obtain estimates for S_r as indicated in Table 1. Dividing the area A of the region by the effective area of each rain gauge S_r, we obtain an estimate of the desirable numbers of gauges needed to reasonably characterize the climatological geostatistical structures of the prevailing rainfall fields across Pakistan.

This analysis suggests that mountainous Northern Pakistan would require on the order of 92-368 quasi-uniformly distributed rain gauges, Central and Southern Pakistan would require on the order of 31-127 and 29-116, respectively. Comparing these numbers to the existing number of gauges in each region provides valuable information regarding how to allocate future resources to improve the hydrometeorological network. For example, the mountainous areas of the Northern Pakistan, which are currently the most under-represented by the existing gauging network, would require at least 10 times the existing rain gauge density to become nominally representative. On the other hand, in the Middle (Southern) regions only 4-5 times increase of the existing rain gauge densities seems to be required. Of course, at sub-seasonal scales, or for tracking individual events, one would require even higher gauge densities, however the estimated gauge densities can be expected to enable the network to provide more reliable information for hydrometeorological applications.

5. CONCLUSIONS

Various climate related applications require spatio-temporally gridded data. Here, we report on the development of a gridded 0.05° latitude-longitude daily data product, PAK-HYM-1.0, that covers the period 1960-2013 over Pakistan. PAK-HYM-1.0 was developed using data from 67 observing stations operated by the National Hydro-Meteorological Service (NHMS) of

Pakistan. It provides gridded fields for daily maximum and minimum temperatures, and accumulated daily precipitation. Compared to other products currently available for the region, PAK-HYM-1.0 has the finest spatial resolution (0.05°) and provides daily data for a longer period (54 years). The closest comparable product in terms of temporal resolution, APHRODITE, is developed using a similar approach but uses a lower gauge density and is not available beyond 2007. The PAK-HYM-1.0 data product discussed here is available from the first author by request, and the original data can be accessed as indicated in Section 3.1.

Based on this data set, we investigate the climatology of various regions of Pakistan, and make recommendations regarding desirable rain gauge densities for each region. As data from other observing stations become digitized and are made available, the product will be progressively updated by NHMS to improve spatio-temporal coverage, and to include gridded fields for other hydro-meteorological variables. Daily updating of the product by NHMS will provide support for weather forecasting and climate hazard monitoring. Work is currently ongoing to use this data product to characterize and investigate the hydrometeorological patterns and thresholds associated with climatic events such as floods, drought and heatwaves over Pakistan.

ACKNOWLEDGMENTS, SAMPLES, AND DATA

We thank Afzaal Karori, Director and Ghulam Rasul, Director General, Pakistan Meteorological Department (PMD) (http://www.pmd.gov.pk/)for providing Precipitation and Temperature data. Furthermore, we are thankful to United States Education Commission Pakistan for providing Fulbright scholarship to Furrukh Bashir to conduct this research. Moreover, we acknowledge, all the datasets utilized in this study by citing them.

REFERENCES

Adler RF, Huffman GJ, Chang A, Ferraro R, Xie P-P, Janowiak J, Rudolf B, Schneider U, Curtis S, Bolvin D, Gruber A, Susskind J, Arkin P, Nelkin E. 2003. The Version-2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979–Present). Journal of Hydrometeorology 4(6): 1147–1167. DOI: 10.1175/1525-7541(2003)004<1147:TVGPCP>2.0.CO;2.

Adler RF, Negri AJ. 1988. A Satellite Infrared Technique to Estimate Tropical Convective and Stratiform Rainfall. Journal of Applied Meteorology, 30–51. DOI: 10.1175/1520-0450(1988)027<0030:ASITTE>2.0.CO;2.

Arkin PA, Joyce R, Janowiak JE. 2017. The estimation of global monthly mean rainfall using infrared satellite data : The GOES precipitation index (GPI) The Estimation of Global Monthly Mean Rainfall Using Infrared Satellite Data : The GOES Precipitation Index (GPI). 7257(July). DOI: 10.1080/02757259409532261.

Arkin P, Meisner B. 1987. The relationship between large-scale convective rainfall and cold cloud over the western hemisphere during 1982-84. Monthly Weather Review, 51–74. DOI: 10.1175/1520-0493(1987)115<0051:TRBLSC>2.0.CO;2.

Ashouri H, Hsu KL, Sorooshian S, Braithwaite DK, Knapp KR, Cecil LD, Nelson BR, Prat OP. 2015. PERSIANN-CDR: Daily precipitation climate data record from multisatellite

observations for hydrological and climate studies. Bulletin of the American Meteorological Society 96(1): 69–83. DOI: 10.1175/BAMS-D-13-00068.1.

Barrett EC. 2000. Satellite remote sensing of precipitation : progress and problems. Remote Sensing Hydrology 2000(267): 3–10.

Becker A, Finger P, Meyer-Christoffer A, Rudolf B, Schamm K, Schneider U, Ziese M. 2013. A description of the global land-surface precipitation data products of the Global Precipitation Climatology Centre with sample applications including centennial (trend) analysis from 1901-present. Earth System Science Data 5(1): 71–99. DOI: 10.5194/essd-5-71-2013.

Behrangi A, Imam B, Hsu K, Sorooshian S, Bellerby TJ, Huffman GJ. 2010. REFAME: Rain Estimation Using Forward-Adjusted Advection of Microwave Estimates. Journal of Hydrometeorology 11: 1305–1321. DOI: 10.1175/2010JHM1248.1.

Berne A, Delrieu G, Creutin JD, Obled C. 2004. Temporal and spatial resolution of rainfall measurements required for urban hydrology. Journal of Hydrology 299(3–4): 166–179. DOI: 10.1016/j.jhydrol.2004.08.002.

Chang CP, Wang Z, Ju J, Li T. 2004. On the relationship between western maritime continent monsoon rainfall and ENSO during northern winter. Journal of Climate 17(3): 665–672. DOI: 10.1175/1520-0442(2004)017<0665:OTRBWM>2.0.CO;2.

Chen M, Shi W, Xie P, Silva VBS, Kousky VE, Higgins RW, Janowiak JE. 2008. Assessing objective techniques for gauge-based analyses of global daily precipitation. Journal of Geophysical Research Atmospheres 113(4): 1–13. DOI: 10.1029/2007JD009132.

Chen M, Xie P, Janowiak JE, Arkin PA. 2002. Global Land Precipitation: A 50-yr Monthly Analysis Based on Gauge Observations. Journal of Hydrometeorology 3(3): 249–266. DOI: 10.1175/1525-7541(2002)003<0249:GLPAYM>2.0.CO;2.

Daly C, Halbleib M, Smith JI, Gibson WP, Doggett MK, Taylor GH, Curtis J, Pasteris PP. 2008. Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. International Journal of Climatology 28(15): 2031–2064. DOI: 10.1002/joc.1688.

Dee DP, Uppala SM, Simmons AJ, Berrisford P, Poli P, Kobayashi S, Andrae U, Balmaseda MA, Balsamo G, Bauer P, Bechtold P, Beljaars ACM, van de Berg L, Bidlot J, Bormann N, Delsol C, Dragani R, Fuentes M, Geer AJ, Haimberger L, Healy SB, Hersbach H, Holm E V., Isaksen L, Kallberg P, Kohler M, Matricardi M, Mcnally AP, Monge-Sanz BM, Morcrette JJ, Park BK, Peubey C, de Rosnay P, Tavolato C, Thepaut JN, Vitart F. 2011. The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. Quarterly Journal of the Royal Meteorological Society 137(656): 553–597. DOI: 10.1002/qj.828.

Duchon CE, Biddle CJ. 2010. Undercatch of tipping-bucket gauges in high rain rate events. Adv. Geosci 25: 11–15. DOI: 10.5194/adgeo-25-11-2010.

Duchon CE, Essenberg GR. 2001. Comparative rainfall observations from pit and aboveground rain gauges with and without wind shields. Water Resources Research 37(12): 3253–3263. DOI: 10.1029/2001WR000541.

Harris I, Jones PD, Osborn TJ, Lister DH. 2014. Updated high-resolution grids of monthly climatic observations - the CRU TS3.10 Dataset. International Journal of Climatology 34(3): 623–642. DOI: 10.1002/joc.3711.

Haylock MR, Hofstra N, Klein Tank AMG, Klok EJ, Jones PD, New M. 2008. A European daily high-resolution gridded data set of surface temperature and precipitation for 1950-2006. Journal of Geophysical Research Atmospheres 113(20). DOI: 10.1029/2008JD010201.

Hou AY, Kakar RK, Neeck S, Azarbarzin AA, Kummerow CD, Kojima M, Oki R, Nakamura K, Iguchi T. 2014. The global precipitation measurement mission. Bulletin of the American Meteorological Society 95(5): 701–722. DOI: 10.1175/BAMS-D-13-00164.1.

Huffman GJ, Adler RF, Morrissey MM, Bolvin DT, Curtis S, Joyce R, McGavock B, Susskind J. 2001. Global Precipitation at One-Degree Daily Resolution from Multisatellite Observations. Journal of Hydrometeorology 2(1): 36–50. DOI: 10.1175/1525-7541(2001)002<0036:GPAODD>2.0.CO;2.

Huffman GJ, Bolvin DT, Braithwaite D, Hsu K, Joyce R, Xie P, Yoo S-H. 2014. NASA global precipitation measurement (GPM) integrated multi-satellite retrievals for GPM (IMERG). Algorithm Theoretical Basis Document (ATBD), NASA/GSFC, Greenbelt, MD, USA

Huffman GJ, Bolvin DT, Nelkin EJ, Wolff DB, Adler RF, Gu G, Hong Y, Bowman KP, Stocker EF. 2007. The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales. Journal of Hydrometeorology 8(1): 38–55. DOI: 10.1175/JHM560.1.

Joyce RJ, Janowiak JE, Arkin PA, Xie P. 2004. CMORPH: A Method that Produces Global Precipitation Estimates from Passive Microwave and Infrared Data at High Spatial and Temporal Resolution. Journal of Hydrometeorology 5(3): 487–503. DOI: 10.1175/1525-7541(2004)005<0487:CAMTPG>2.0.CO;2.

Kidd C, Huffman G. 2011. Global precipitation measurement. Meteorological Applications 18(3): 334–353. DOI: 10.1002/met.284.

Kidd C, Kniveton DR, Todd MC, Bellerby TJ, Kidd C, Kniveton DR, Todd MC, Bellerby TJ. 2003. Satellite Rainfall Estimation Using Combined Passive Microwave and Infrared Algorithms. Journal of Hydrometeorology 4(6): 1088–1104. DOI: 10.1175/1525-7541(2003)004<1088:SREUCP>2.0.CO;2.

Kubota T, Hashizume H, Shige S, Okamoto K, Aonashi K, Takahashi N, Ushio T, Kachi M. 2006. Global precipitation map using satelliteborne microwave radiometers by the GSMaP project: Production and validation. International Geoscience and Remote Sensing Symposium (IGARSS), 2584–2587. DOI: 10.1109/IGARSS.2006.668.

Kummerow C, Barnes W, Kozu T, Shiue J, Simpson J. 1998. The Tropical Rainfall Measuring Mission (TRMM) sensor package. Journal of Atmospheric and Oceanic Technology 15(3): 809–817. DOI: 10.1016/0273-1177(94)90210-0.

Lau KM, Wu HT. 2001. Principal modes of rainfall-SST variability of the Asian Summer Monsoon: A reassessment of the Monsoon-ENSO relationship. Journal of Climate 14(13): 2880–2895. DOI: 10.1175/1520-0442(2001)014<2880:PMORSV>2.0.CO;2.

Mariotti A. 2007. How ENSO impacts precipitation in southwest central Asia. Geophysical Research Letters 34(16): 2–6. DOI: 10.1029/2007GL030078.

Nešpor V, Krajewski WF, Kruger A. 2000. Wind-Induced Error of Raindrop Size Distribution Measurement Using a Two-Dimensional Video Disdrometer. Journal of Atmospheric and Oceanic Technology 17(11): 1483–1492. DOI: 10.1175/1520-0426(2000)017<1483:WIEORS>2.0.CO;2.

Osborn TJ, Hulme M. 1998. Evaluation of the European daily precipitation characteristics from the atmospheric model intercomparison project. International Journal of Climatology 18(5): 505–522. DOI: 10.1002/(SICI)1097-0088(199804)18:5<505::AID-JOC263>3.0.CO;2-7.

Reinert D, Prill F, Frank H, Zängl G. 2016. A global satellite-assisted precipitation climatology. Deutscher Wetterdienst, Research and Development (FE1 3) (1994): 2000. DOI: 10.5676/DWD.

Schneider U, Fuchs T, Meyer-Christoffer A, Rudolf B. 2011. Global Precipitation Analysis Products of the GPCC. Global Precipitation Climatology Centre (GPCC). Gpcc.Dwd.De 12.

Scofield RA, Kuligowski RJ. 2003. Status and Outlook of Operational Satellite Precipitation Algorithms for Extreme-Precipitation Events. Weather and Forecasting 18(6): 1037–1051. DOI: 10.1175/1520-0434(2003)018<1037:SAOOOS>2.0.CO;2.

Shepard D. 1968. A two-dimensional interpolation function for irregularly-spaced data. Proceedings of the 1968 23rd ACM national conference on - 517–524. DOI: 10.1145/800186.810616.

Sorooshian S, Hsu KL, Gao X, Gupta H V., Imam B, Braithwaite D. 2000. Evaluation of PERSIANN system satellite-based estimates of tropical rainfall. Bulletin of the American Meteorological Society 81(9): 2035–2046. DOI: 10.1175/1520-0477(2000)081<2035:EOPSSE>2.3.CO;2.

Stahl K, Moore RD, Floyer JA, Asplin MG, McKendry IG. 2006. Comparison of approaches for spatial interpolation of daily air temperature in a large region with complex topography

and highly variable station density. Agricultural and Forest Meteorology 139(3–4): 224–236. DOI: 10.1016/j.agrformet.2006.07.004.

Stephens GL, Vane DG, Boain RJ, Mace GG, Sassen K, Wang Z, Illingworth AJ, O'Connor EJ, Rossow WB, Durden SL, Miller SD, Austin RT, Benedetti A, Mitrescu C. 2002. The cloudsat mission and the A-Train: A new dimension of space-based observations of clouds and precipitation. Bulletin of the American Meteorological Society 83(12): 1771–1790+1742. DOI: 10.1175/BAMS-83-12-1771.

Tapiador FJ, Navarro A, Levizzani V, García-Ortega E, Huffman GJ, Kidd C, Kucera PA, Kummerow CD, Masunaga H, Petersen WA, Roca R, Sánchez J-L, Tao W-K, Turk FJ. 2017. Global precipitation measurements for validating climate models. Atmospheric Research 197(May): 1–20. DOI: 10.1016/j.atmosres.2017.06.021.

Tapiador FJ, Turk FJ, Petersen W, Hou AY, García-Ortega E, Machado LAT, Angelis CF, Salio P, Kidd C, Huffman GJ, de Castro M. 2012. Global precipitation measurement: Methods, datasets and applications. Atmospheric Research. Elsevier B.V. 104–105: 70–97. DOI: 10.1016/j.atmosres.2011.10.021.

Taylor KE. 2001. in a Single Diagram. Journal of Geophysical Research 106(D7): 7183–7192. DOI: 10.1029/2000JD900719.

Teegavarapu RS V, Chandramouli V. 2005. Improved weighting methods, deterministic and stochastic data-driven models for estimation of missing precipitation records. Journal of Hydrology 312(1–4): 191–206. DOI: 10.1016/j.jhydrol.2005.02.015.

Turk F. J, Hawkins J, Smith EA, Marzano FS, Mugnai A, Levizzani V. 2000. Combining SSM/I, TRMM and Infrared Geostationary Satellite Data in a Near-realtime Fashion for Rapid Precipitation Updates: Advantages and Limitations. The 2000 EUMETSAT Meteorological Satellite Data Users' Conference 452–459.

Vicente GA, Scofield RA, Menzel WP. 1998. The 0perational GOES infrared rainfall estimation technique. Bulletin of the American Meteorological Society 79(9): 1883–1898. DOI: 10.1175/1520-0477(1998)079<1883:TOGIRE>2.0.CO;2.

Viney NR, Bates BC. 2004. It never rains on Sunday: the prevalence and implications of untagged multi-day rainfall accumulations in the Australian high quality data set. International Journal of Climatology 24(9): 1171--1192.

Willmott CJ, Rowe CM, Philpot WD. 1985. Small-Scale Climate Maps: A Sensitivity Analysis of Some Common Assumptions Associated with Grid-Point Interpolation and Contouring. Cartography and Geographic Information Science 12(1): 5–16. DOI: 10.1559/152304085783914686.

Xie P, Arkin PA. 1997. Global Precipitation: A 17-Year Monthly Analysis Based on Gauge Observations, Satellite Estimates, and Numerical Model Outputs. Bulletin of the American Meteorological Society 78(11): 2539–2558. DOI: 10.1175/1520-0477(1997)078<2539:GPAYMA>2.0.CO;2.

Xie P, Chen M, Yang S, Yatagai A, Hayasaka T, Fukushima Y, Liu C. 2007. A Gauge-Based Analysis of Daily Precipitation over East Asia. Journal of Hydrometeorology 8(3): 607–626. DOI: 10.1175/JHM583.1.

Xie P, Janowiak JE, Arkin PA, Adler R, Gruber A, Ferraro R, Huffman GJ, Curtis S. 2003. GPCP pentad precipitation analyses: An experimental dataset based on gauge observations and satellite estimates. Journal of Climate 16(13): 2197–2214. DOI: 10.1175/2769.1.

Yatagai A, Kamiguchi K, Arakawa O, Hamada A, Yasutomi N, Kitoh A. 2012. Aphrodite constructing a long-term daily gridded precipitation dataset for Asia based on a dense network of rain gauges. Bulletin of the American Meteorological Society 93(9): 1401–1415. DOI: 10.1175/BAMS-D-11-00122.1.

Table B. 1: Estimates of desirable numbers of gauges for different parts of Pakistan, based on geo-statistical analysis of climatological summer precipitation (See Figure B. 7).

	Area (%)	Decorrelation Distance r (km)	Representative Surface Area S_r of a Gauge $\left(\Delta = \frac{1}{2}r\right)$ (km^2)	Desirable # of Gauges (min-max) ($\sim Area/S_r$)	Existing # of Gauges
Area > 35°N	11.14	75	481	92-368	10
30°N < Area < 35°N	34.58	210	4329	31-127	31
Area < 30°N	54.28	275	7424	29-116	26
			Total Gauges Required	152-612	

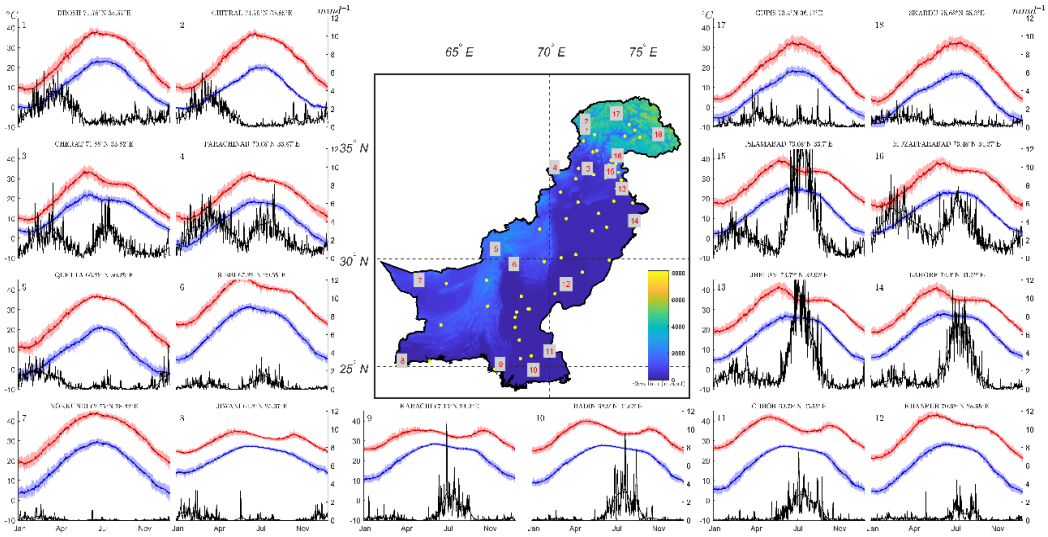


Figure B. 1. Hypsometry and climatology of Pakistan: Map presenting elevation of Pakistan (Center) along with 365-day climatologies of maximum temperature (red line), minimum temperature (blue line) and precipitation (black line) at various locations of the country (subplots 1-12). Each bold line indicates the 365-day smoothed climatology for a specific location and the filled region indicates the 25th to 75th percentile spread. The y-axis indicates temperature ($°C$) on odd numbered subplots and precipitation (mm/day) on even numbered subplots.

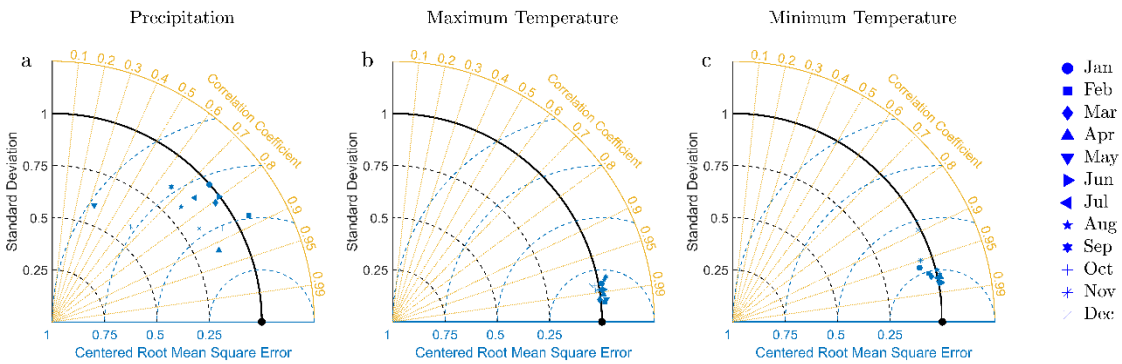


Figure B. 2. Taylor Diagrams presenting spatial cross-validation of ICDWM imputation in time series for year 2011 via the leave-one-out method; (a) Mean Precipitation, (b) Mean Maximum Temperature and (c) Mean Minimum Temperature, all at monthly scales. Each plot presents the centered root mean square error (CRMSE) on the x-axis, normalized standard deviation on the y-axis, and correlation between observations and reconstructed time series along the circumference of the quarter circle. See section 3.3 for details.

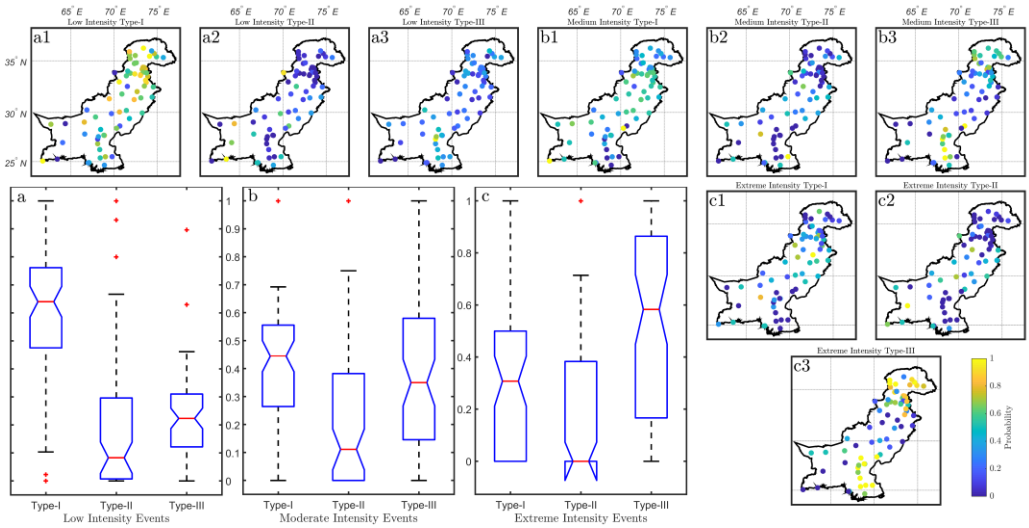


Figure B. 3. Bar and whisker plots indicating the probability of type I, type II, and type III events for (a) low intensity CAT-I, (b) moderate intensity CAT-II, and (c) extreme intensity CAT-III events. Subplots (a1-a3, b1-b3, c1-c3) indicate, at observations sites, the probabilities of event types in each category. See section 4.1 for details.

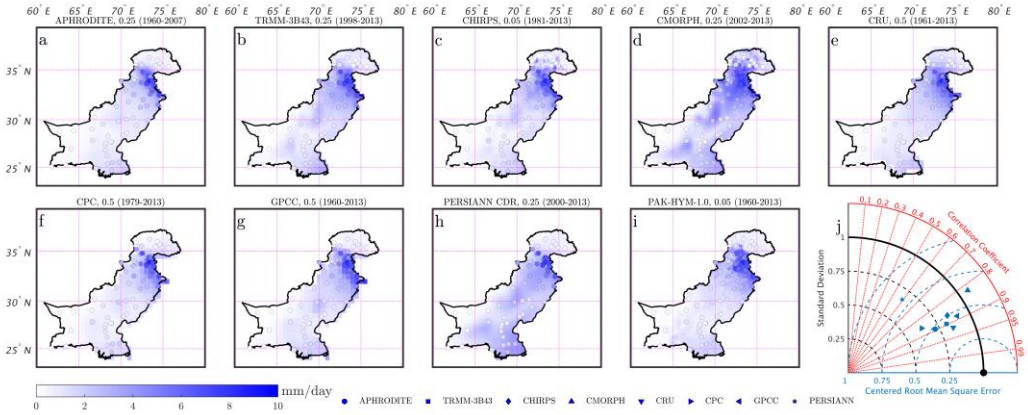


Figure B. 4. Comparative evaluation of PAK-HYM-1.0 mean summer precipitation (mm/day) with alternative precipitation products; (a) APHRODITE (1960-2007), (b) TRMM-3b43 (1998-2013), (c) CHIRPS (1981-2013), (d) CMORPH (2002-2013), (e) CRU (1961-2013), (f) CPC (1979-2013), (g) GPCC (1960-2013), (h) PERSIANN (2000-2013), (i) PAK-HYM-1.0 (1960-2013). Subplot (j) presents a spatial statistical comparison of PAK-HYM-1.0 with other products for the summer season. See section 4.2 for details.

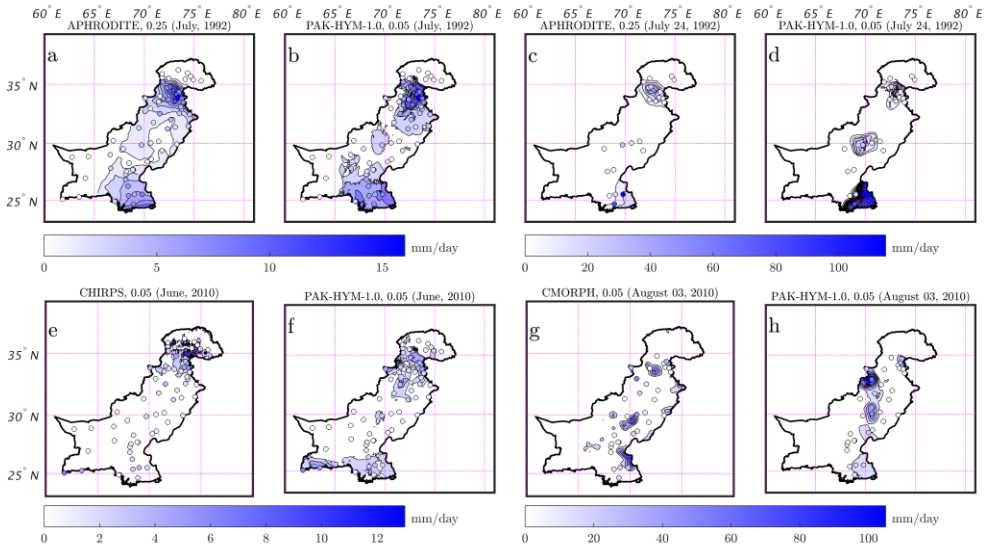


Figure B. 5. Comparison of PAK-HYM-1.0 with APHRODITE, CHIRPS and CMORPH precipitation products for individual events at the monthly and daily scales; (a & b) Monthly comparison with APHRODITE for July 1992, (c & d) Daily comparison with APHRODITE for July 24, 1992, (e & f) Monthly comparison with CHIRPS for June 2010, (g & h) Daily comparison with CMORPH for Aug 03, 2010. See section 4.2 for details.

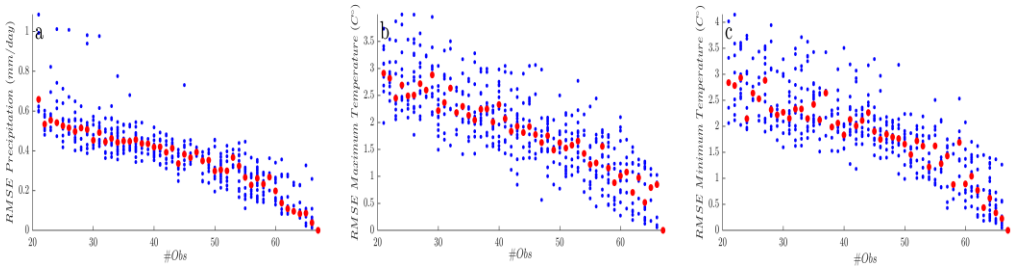


Figure B. 6. Sensitivity of the interpolation algorithm to number of observations used. The y-axis shows RMSE for multiple realizations and x-axis shows observational density; (a) Mean annual precipitation, (b) Mean annual maximum temperature, (c) Mean annual minimum temperature. See section 4.3 for details.

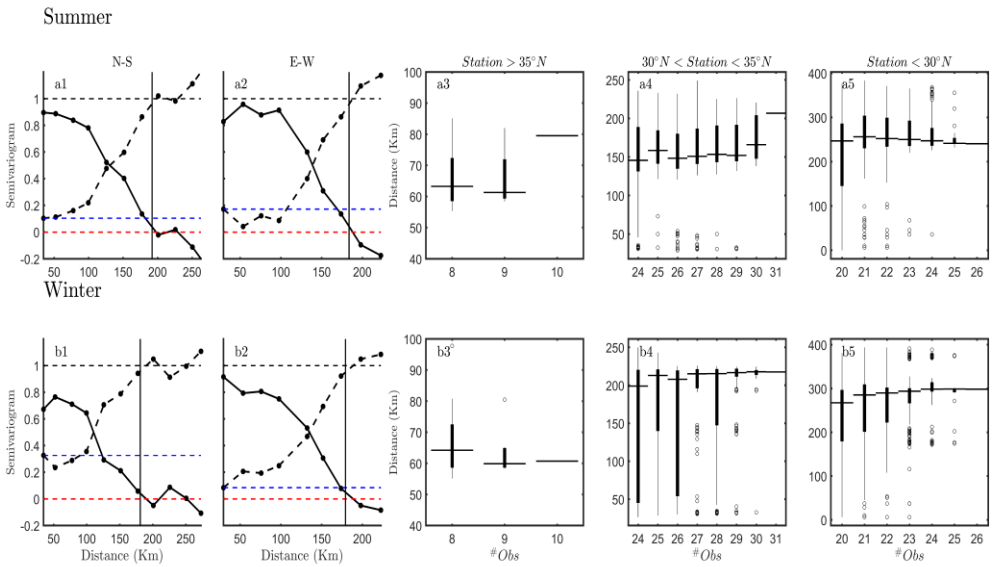


Figure B. 7. The first two columns show experimental semi-variograms for mean summer and winter precipitation in Pakistan with directional anisotropy; (a1) Summer N-S, (a2) Summer E-W, (b1) Winter N-S, (b2) Winter E-W. The dashed and solid black lines indicate the semi-variance (γ) and covariance (C), respectively. The dotted red, blue and black lines indicate 0, the nugget and the sill, respectively. The last three columns show observational site decorrelation lengths for different numbers of sites; (a3) summer above 35°N, (b3) winter above 35°N, (a4) summer between 30°N and 35°N, (b4) winter between 30°N and 35°N, (a5) summer below 30°N, (b5) winter below 30°N. The bar and whisker plots indicate uncertainties associated with the de-correlation lengths (y-axis) estimated from random subsets of the observations. See section 3.5 and 4.3 for details.

SUPPORTING INFORMATION

A BRIEF REVIEW OF MAJOR SOURCES OF HYDRO-METEOROLOGICAL DATA DEVELOPMENT

High quality temperature and precipitation dataset is a primary requirement of the resource management and hydro-climatic monitoring studies. Among hydro-meteorological variables, precipitation is one of the most challenging quantities to estimate, because of its high spatial and temporal variability.

Presently, major sources of temperature and precipitation data generation are in-situ observations, satellite retrieved information and outputs from numerical models, however, each has its own merits and demerits (Xie and Arkin, 1997).

First, we evaluate the efficiency of the in-situ rain gauge observations. In-situ gauge collecting precipitation is the most rudimentary, common, direct, longest-lived precipitation quantification method. With the technological advancements, the simple collecting vessels, which are considered as standard accumulation gauges, are upgraded to sophisticated tipping bucket gauges, weighing gauges, optical gauges and disdrometers; all have their own characteristics with relative advantages and disadvantages. At gauge scale, the catch of precipitation by a gauge is affected by the wind flow around the orifice (Duchon and Essenberg, 2001). Wind-gauge interaction interrupts the flow across the gauge orifice affecting light precipitation the most and heavy downpour to some extent (Duchon and Biddle, 2010), resulting in under-catch at low intensities and higher windspeeds (Nešpor *et al.*, 2000). Moreover, rain gauge observations suffer from different types of rain gauges operated by different environmental surveillance agencies and their reporting regulations.

Another important aspect in collection of observation from rain gauges is the timing of the observations, which differs between different National Hydro-meteorological Services (NHMS). This sometimes creates untagged multi-day accumulations with clear implications for the comparisons with the sources of data (Viney and Bates, 2004). While gauge data are generally used “as is”, errors and uncertainties associated with such precipitation measurements are reasonably well understood and corrections or further quality control measures can be applied.

However, despite errors, rain gauges remain arguably the most accurate instrument to measure rainfall at the surface, and conventional observations through surface gauge network provide the most direct measure of precipitation (Kidd and Huffman, 2011). Estimates of air temperature from station network have similar causes of uncertainty as gauge-based precipitation estimates; they lack sufficient station density to resolve the spatial variability of temperature, particularly in complex terrain [e.g., *Stahl et al.*, 2006].

Second major source of precipitation estimation is satellite observations. Presently, history of the satellite retrieved precipitation and temperature is nearly four-decade old and it covers nearly whole globe. Satellite retrieved precipitation and temperature products are superior to their ground observed counterparts in regard to their spatial coverage but are subject to retrieval errors and biases. Satellite observations from visible, infrared, passive and active microwave systems are utilized to render precipitation estimates using a number of techniques that are different in performance on regional and temporal scales (Kidd and Huffman, 2011). For precipitation estimation, observations in the thermal infrared channels allow cloud top temperatures to be retrieved that are further used to retrieve precipitation rate.

However, the actual relationship is considerably more complicated than this simple interpretation of cloud top temperature into precipitation. Several infrared based techniques are providing precipitation information on different spatial and temporal resolutions, such as estimation of convective precipitation derived from geostationary satellite imagery using a simple thresholding technique (Arkin and Meisner, 1987), estimation of both tropical convective precipitation and stratiform precipitation from satellite infrared data (Adler and Negri, 1988), estimation of precipitation rate through the Auto-Estimator (Vicente *et al.*, 1998), the quantitative precipitation estimator (QPE) for extreme events (Scofield and Kuligowski, 2003), and the Precipitation Estimation from Remote-Sensed Information using Artificial Neural Network [PERSIANN*; Sorooshian et al.*, 2000] that uses various satellite and ground based information sources to estimate precipitation.

Furthermore, microwave (MW) based methods that utilizes upwelling microwave energy to explain the distribution of precipitation-sized hydrometeors includes the Tropical Rainfall

Measuring Mission [TRMM; *Kummerow et al.*, 1998], the CloudSat; which measure cloud vertical structure and precipitation and carries a cloud radar aboard (Stephens *et al.*, 2002), and the Global Precipitation Measurement (GPM) mission (Hou *et al.*, 2014) which is comprised of a constellation of microwave sensors. However, their coverage may be limited to specific latitude band and estimates from any single product are infrequent that provides nothing more than a single snapshot (Tapiador *et al.*, 2017). On the other hand, the relative high frequency sampling of geostationary (visible/IR channels) satellites is compromised by the indirectness of the observations.

However, both IR and MW can be combined to get a better precipitation estimating product. One way is to calibrate IR observations by passive microwave (PMW) estimates as The NRL-Blended technique (Turk *et al.*, 2000) and the Passive Microwave-InfraRed (PMIR) technique (Kidd *et al.*, 2003) uses a moving spatial and temporal window to generate a local relationship between the fast-refresh IR observations and the precipitation estimates sourced from the Level-2 (i.e., instantaneous swath-level) PMW observations (Tapiador *et al.*, 2017). In another approach, the precipitation estimates from relatively infrequent PMW (polar orbiting satellites) are integrated with geostationary satellite based IR data to track the trajectory of the precipitation between adjacent PMW estimates. These are called advection morphed or Lagrangian time-interpolation schemes. Climate Prediction Center (CPC) Morphing [CMORPH; *Joyce et al.*, 2004], and the Global Satellite Map Product [GSMaP; *Kubota et al.*, 2006], and the Integrated Multi-satellitE Retrievals for GPM [IMERG; *Huffman et al.*, 2014], and the Rain Estimation using Forward Adjusted-advection of Microwave Estimates [REFAME; *Behrangi et al.*, 2010]. All these stated products use ground observation for calibration or verification for their efficiency.

Moreover, satellite based precipitation estimates, sometimes, incorporates surface precipitation observations as an anchor point to generate a high spatiotemporal resolution product. Such products can be classified into two broad classes: i.e., Algorithms that emphasize homogeneous inputs and processing, are referred to as Climate Data Records (CDR), and those that gives the best short-interval estimates are referred to as High-Resolution Precipitation

Products [HRPP; *Tapiador et al.*, 2017]. The TRMM Multi-Satellite Precipitation Analysis [TMPA; *Huffman et al.*, 2007] a global (50ºN-50ºS) 3-hourly 0.25º×0.25º resolution data from 1998-present is an example of HRPP.

Among CDRs, various products are presented by Global Precipitation Climatology Project (GPCP) that includes; a global 2.5º×2.5º merged Satellite-Gauge (SG) precipitation analysis estimates on monthly scale from 1979 to the present (Adler *et al.*, 2003), One-Degree Daily (1DD) precipitation analysis [October 1966 – onward; *Huffman et al.*, 2001] that uses a Threshold-Matched Precipitation Index (TMPI) to estimate instantaneous precipitation from geostationary Infrared (GEO-IR), a 2.5º×2.5º monthly global Merged Analysis of Precipitation (CMAP) from rain gauges and satellite precipitation estimates (Xie and Arkin, 1997), and another GPCP 2.5º×2.5º pentad precipitation analysis that uses the SG product to adjust the pentad CPC Merged Analysis of Precipitation (CMAP) pentad product so that the overall magnitude of the product matches at the monthly scale but with the sub-monthly variability of pentad CMAP product (Tapiador *et al.*, 2017; Xie *et al.*, 2003). It is worth mentioning that most of the merged products, either CDR or HRPP, were never intended to be used for the trends estimation as their input data sources drift.

Similarly, air temperature estimates from satellite are limited due to intrinsic characteristics of the satellite and availability of limited in-situ meteorological data.

Third source is Numerical Weather Prediction (NWP) models and Retrospective-analysis (reanalysis) that also provide precipitation estimates. However, mostly NWP and reanalysis estimate precipitation in forecast mode based on temperature and precipitation information derived from assimilated observations. Such methodology may under/overestimate the precipitation that is corrected by gridded observed precipitation products (Dee *et al.*, 2011). Moreover, the efficiency of reanalysis precipitation estimation is further evaluated by comparing it with gridded observed products. Hence, the gridded observations are usually considered as the first approximation of the truth.

Gridded observed products are comprised of historical precipitation records mainly collected at land surface (irregular meteorological observatories) presented in a gridded format.

Such data sets that are interpolated at regular grids allow best estimates of climatic variables at locations away from the observation collection sites and thus provide information regarding variable under consideration at data-sparse locations (Haylock *et al.*, 2008). The original data in the different gridded data sets is the same sometimes, however, the gridded data may differ in how the station records are filtered, interpolated, and homogenized so the agreement among different gridded data products is not surprising (Tapiador *et al.*, 2012). Most of the gridded precipitation and temperature products consider topographic effects, that may introduce a large error in the results due to disparity of elevation model from original hypsometric heights. Never the less, areas where observation density is higher, such topographic errors may be masked, as information density may overcome the discrepancies of the elevation model.

The Climate Research Unit (CRU, https://crudata.uea.ac.uk/cru/data/hrg/) at the University of East Anglia has developed one of the leading gridded observed data sets that includes absolute monthly values of precipitation, mean temperature with diurnal temperature, vapor pressure, along with cloud cover and wet-day frequency at 0.5º×0.5º global grid covering period of 1900-2013 (Harris *et al.*, 2014). Another leading gridded data set is developed by Global Precipitation Climatology Center (GPCC, http://www.dwd.de/EN/ourservices/gpcc/gpcc.html) at German Weather Service (DWD) under auspices of World Meteorological Organization (WMO). GPCC presents a number of products and claim that its archive is by far the largest worldwide for monthly precipitation, outperforming the global precipitation data coverage of other similar data sets by at least a factor of two and partly much more (Becker *et al.*, 2013). Its finest resolution products available is 0.25º ×0.25º that presents climatology (1951-2000).

Furthermore, the US Climate Prediction Center (CPC, ftp://ftp.cpc.ncep.noaa.gov/precip/50yr/gauge/50yr/gauge/0.5deg/) produces, the Monthly Analysis of Global Land Precipitation from 1948 to the Present, at a 0.5°×0.5° spatial resolution [PREC/L; *Chen et al.*, 2002]. At regional scale, another gridded data set Asian Precipitation-Highly Resolved Observational Data Integration Towards Evaluation of Water Resources (APHRODITE, http://search.diasjp.net/en/dataset/APHRO_ME) at 0.25° and 0.5° resolution daily precipitation dataset presented by a consortium between the Research Institute for Humanity and Nature (RIHN) Japan

and the Meteorological Research Institute of Japan Meteorological Agency (MRI/JMA) is developed from dense rain gauge network. It covers 60°E–150°E, 15°S–55°N for the period extending from 1951–2007 (Yatagai *et al.*, 2012).

DIGITAL ELEVATION MODEL

Relationship of precipitation with elevation is highly variable, nevertheless, generally it increases with elevation, with the exception of terrain above moist boundary layer (Daly *et al.*, 2008), or exposure of a slope to moisture laden winds. Whereas, temperature has strong reliable relation to elevation, that it usually decreases with the increase in height. Exceptions, such as inversion in cold regions or in winter season, or summer marine layers in coastal areas, where maximum temperatures often increase with elevation above the marine inversion do exists. Nevertheless, for a localized region elevation is the most important factor in the distribution of temperature and precipitation. Structure of land's hypsometry is best described by a digital elevation model (DEM).

DEM is a matrix of cells, covering some part of the earth, while maintaining spatial continuity, each cell represents a unit square area and its average elevation. Agreement between in-situ measured elevation and DEM is better in plain as compared to complex and highly erratic topography. Using DEM, a local climate-elevation relationship for each grid cell is developed through regression, where slope of the line is dictated by the elevation and magnitude of the participating points. If elevation of all participating sites and target location do not have a very large deviation from each other, such regression provides better results. Whereas, if target grid cell's elevation is way below or above the mean elevation of the neighboring participating sites, the linear regression may estimate erroneous values that do not match with the climatology of the target cell. Here, we use ASTER (30m) DEM, developed by NASA (http://www.nasa.network.com/elev?) to obtain the elevation information about the terrain of Pakistan.

INTERPOLATION SCHEME

Based on a simple hypothesis that the value (noise) of each nearby point influences the estimation of the associated grid-point value (noise), The weights are ascribed to three categories (Willmott *et al.*, 1985):

$$S_k = \begin{cases} d_{i,k}^{-1}, & d_{j,k} \leq \frac{r_j}{3} \\ \frac{27}{4r_j}\left(\frac{d_{j,k}}{r_j} - 1\right)^3, & \frac{r_j}{3} < d_{j,k} \leq r_j \\ 0, & d_{j,k} > r_j \end{cases} \quad \text{(s1)}$$

where r_j is the search radius and $d_{j,k}$ is the distance from grid point j to nearby data point k. The subscript k is used to refer one of those data points near j that belong to the set of data points influencing the interpolated value at j. The research radius is defined as a constant radius of a circle that is equal to an average of 10 data points. On the sphere, the great-circle distance between points j and k is obtained from

$$\cos d_{j,k}^s = sin\phi_j sin\phi_k + cos\phi_j cos\phi_k(\lambda_j - \lambda_k) \quad \text{(s2)}$$

$$d_{j,k}^c = \left[(x_k - x_j)^2 + (y_k - y_j)^2\right]^{0.5} \quad \text{(s3)}$$

where superscripts c and s, distinguish cartesian from spherically derived distances. Once weights ($S_k = 1, n_j$) have been estimated, this can be further corrected for the "directional isolation" of each data point relative to all other nearby points through

$$T_k = \sum_{l=1}^{n_j} S_l[1 - cos\theta_j(k,l)], j \neq k \quad \text{(s4)}$$

where $\theta_j(k,l)$ is the angular separation of nearby data points k and l when the vertex of the angle is defined as grid-point j. The angular solitude of a data point with respect to j, on the sphere is

$$cos\theta_j^s(k,l) = \frac{cosd_{k,l}^s - cosd_{j,k}^s\, cosd_{j,l}^s}{sind_{j,k}^s sind_{j,l}^s} \quad \text{(s5)}$$

Therefore, the data points having small angular separation contribute less, individually, than points having a large angular separation. Now, with the directional isolation of data-point k known, a combined set of weights is calculated from

$$W_k = S_k^2\left(1 + \frac{T_k}{\sum_{l=1}^{n_j} S_l}\right),\ l \neq k, \qquad \text{(s6)}$$

which limits the maximum influence of the directional isolation of k to twice the weights based on distance, that is when

$$T_k = \sum\nolimits_{l=1}^{n_j} S_l \qquad \text{(s7)}$$

To obtain the non-zero gradients on the interpolated surface at data points, increments (Δz_k) are computed and added to the respective data-point values (z_k). The correction involves finding as average weighted gradient for each of the n_j points within r_j of j, based upon the collective rates of change at the other data points within r_j. On the surface of the sphere, the incremental corrections are obtained for each of the data points associated with j from

$$\Delta z_k = \left\{ \overline{\frac{\partial(\Delta z)}{\partial\lambda}}\Bigg]_k d_\lambda^s(j,k) + \overline{\frac{\partial(\Delta z)}{\partial\phi}}\Bigg]_k d_\phi^s(j,k)\right\} \times \{v/(v + d_{j,k}^s)\}, \qquad \text{(s8)}$$

where $\overline{\frac{\partial(\Delta z)}{\partial\lambda}}$ is the average partial derivative with respect to longitude, $\overline{\frac{\partial(\Delta z)}{\partial\phi}}$ is the average partial derivative with respect to latitude, $d_\lambda^s(j,k) = (\lambda_j - \lambda_k)\cos\phi$, and $d_\phi^s(j,k) = (\phi_j - \phi_k)$. The average partial derivatives are taken as:

$$\overline{\frac{\partial(\Delta z)}{\partial\lambda}}\Bigg]_k = \frac{\sum_{l=1}^{n_j} W_l(z_l - z_k)d_\lambda^s(l,k)d_{k,l}^{s-2}}{\sum_{l=1}^{n_j} W_l},\ l \neq k, \qquad \text{(s9)}$$

and

$$\left.\overline{\frac{\partial(\Delta z)}{\partial\phi}}\right]_k = \frac{\sum_{l=1}^{n_j} W_l(z_l - z_k) d_\phi^s(l,k) d_{k,i}^{s-2}}{\sum_{l=1}^{n_j} W_l},\ l \neq k, \qquad \text{(s10)}$$

where $d_\lambda^s(l,k) = (\lambda_l - \lambda_k)\cos\phi$ and $d_\phi^s(l,k) = (\phi_l - \phi_k)$, whereas a somewhat arbitrary adjustment is made to limit the influence of the increments to one-tenth the data range, so $|\Delta z_k|$ cannot exceed 0.1(max z_i – min z_i), where i again refers to any point on the map. The adjustment parameter (v) takes the form

$$v = \frac{0.1(\max\{z_i - \min z_i)\}}{max\left[\left(\left.\overline{\frac{\partial(\Delta z)}{\partial\lambda}}\right]_k\right)^2 + \left(\left.\overline{\frac{\partial(\Delta z)}{\partial\phi}}\right]_k\right)^2\right]^{0.5}} \qquad \text{(s11)}$$

With the evaluation of these functions, the value predicted at grid point $j(\widehat{z_j})$ becomes

$$\widehat{z_j} = \frac{\sum_{k=1}^{n_j} W_k(z_k + \Delta z_k)}{\sum_{k=1}^{n_j} W_k} \qquad \text{(s12)}$$

where, m is the number of data points with in the search radius.

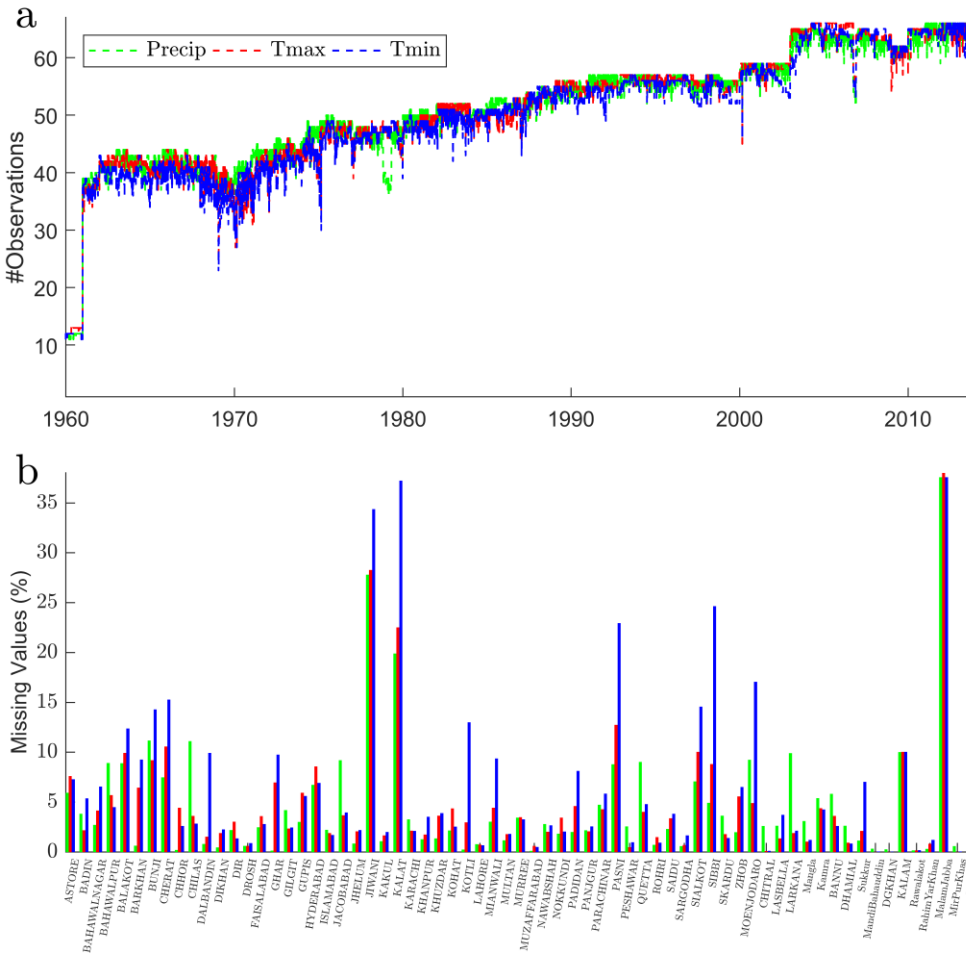


Figure B.S1. (a) Timeseries of available observations of precipitation, maximum and minimum temperature from 1960 to 2013, (b) percentage of missing values relative to total observations for each site.

Table B S1. Description of gridded observed and satellite products used for comparison.

Product Name	Temporal Resolution	Spatial Resolution	Duration	Producer (Developer)
Aphrodite	Daily	0.25°	1951-2007	(Yatagai *et al.,* 2012) http://search.diasjp.net/en/dataset/APHRO_ME
Climate Prediction Center(CPC)	Monthly	0.5°	1948 to current	Chen et al., 2002 ftp://ftp.cpc.ncep.noaa.gov/precip/50yr/gauge/50yr/gauge/0.5deg/
CRU TS 310	Monthly	0.5°	1901-2015	Harris at al 2014 https://crudata.uea.ac.uk/cru/data/hrg/
GPCC	Monthly	0.5°	1901-2013	(Schneider *et al.*, 2011) http://www.dwd.de/EN/ourservices/gpcc/gpcc.html
TRMM-3B43	Monthly	0.25°	1998-2015	(Huffman *et al.*, 2007) https://pmm.nasa.gov/data-access/downloads/TRMM
CMORPH	Daily	0.25°	2002-Present	(Joyce et al., 2004) ftp://ftp.cpc.ncep.noaa.gov/precip/global_CMORPH
PERSIANN CDR	Daily	0.25°	1983-Present	(Ashouri et al., 2015) http://chrsdata.eng.uci.edu/
CHIRPS	Monthly	0.05°	1981- Present	(Reinert et al., 2016) http://chirp.danplanet.com/projects/chirp/wiki/Download

Table B.S2. Name of the observational sites, their location and elevation, along with percentage of missing data.

Station Name	Longitude (°E)	Latitude (°N)	Elevation (m.a.s.l)	From	To	Precipitation Missing Values (%)	Maximum Temperature Missing Values (%)	Minimum Temperature Missing Values (%)
ASTORE	74.90	35.37	2167	01-Jan-1961	31-Dec-2013	5.9	7.6	7.31
BADIN	68.90	24.63	10	01-Jan-1961	31-Dec-2013	3.8	2.2	5.40
BAHAWALNAGAR	73.25	29.95	161	01-Aug-1962	31-Dec-2013	2.7	4.2	6.58
BAHAWALPUR	71.78	29.40	116	01-Jan-1961	31-Dec-2013	8.9	5.7	4.51
BALAKOT	73.35	34.38	980	01-Jul-1961	31-Dec-2013	8.9	9.9	12.39
BARKHAN	69.72	29.88	1097	01-Jan-1974	31-Dec-2013	0.6	6.5	9.29
BUNJI	74.63	35.67	1372	01-Jan-1961	31-Dec-2013	11.2	9.2	14.31
CHERAT	71.88	33.82	1301	01-Feb-1961	31-Dec-2013	7.5	10.6	15.29
CHHOR	69.78	25.52	5	01-Jan-1961	31-Dec-2013	0.2	4.4	2.62
CHILAS	74.10	35.42	1250	01-Jan-1961	31-Dec-2013	11.1	3.6	2.85
DALBANDIN	**64.40**	**28.88**	**848**	**01-Jan-1961**	**31-Dec-2013**	**0.8**	**1.5**	**9.95**
DERA ISMAEEL KHAN	**70.92**	**31.82**	**173**	**01-Jan-1960**	**31-Dec-2013**	**0.5**	**1.9**	**2.28**
DIR	**71.85**	**35.20**	**1369**	**01-Jan-1968**	**31-Dec-2013**	**2.2**	**3.1**	**1.38**
DROSH	**71.78**	**35.57**	**1464**	**01-Jan-1961**	**31-Dec-2013**	**0.6**	**0.6**	**0.91**
FAISALABAD	**73.10**	**31.43**	**183**	**01-Jan-1960**	**31-Dec-2013**	**2.5**	**3.6**	**2.82**
GHARI DUPATTA	**73.62**	**34.22**	**812**	**01-Jan-1961**	**31-Dec-2013**	**0.2**	**7.0**	**9.76**
GILGIT	**74.33**	**35.92**	**1459**	**01-Jan-1960**	**31-Dec-2013**	**4.2**	**2.4**	**2.47**
GUPIS	**73.40**	**36.17**	**2155**	**01-Jan-1961**	**31-Dec-2013**	**3.0**	**5.9**	**5.62**
HYDERABAD	**68.42**	**25.38**	**40**	**01-Jan-1961**	**31-Dec-2013**	**6.7**	**8.6**	**6.94**
ISLAMABAD	**73.08**	**33.70**	**543**	**01-Jan-1960**	**31-Dec-2013**	**2.2**	**1.9**	**1.73**
JACOBABAD	**68.47**	**28.30**	**55**	**01-Jan-1961**	**31-Dec-2013**	**9.2**	**3.7**	**3.97**
JHELUM	**73.72**	**32.93**	**232**	**01-Jan-1960**	**31-Dec-2013**	**0.8**	**2.1**	**2.22**
JIWANI	**61.80**	**25.07**	**56**	**01-Jan-1961**	**31-Dec-2013**	**27.8**	**28.3**	**34.38**
KAKUL	**73.25**	**34.18**	**1308**	**01-Jan-1960**	**31-Dec-2013**	**1.1**	**1.7**	**2.02**
KALAT	66.58	29.03	2015	01-Jan-1961	31-Dec-2013	19.9	22.5	37.27
KARACHI	67.13	24.90	21	01-Jan-1961	31-Dec-2013	3.3	2.2	2.13
KHANPUR	70.68	28.65	87	01-Jan-1961	31-Dec-2013	1.3	1.8	3.55
KHUZDAR	66.63	27.83	1231	01-Jan-1975	31-Dec-2013	1.4	3.7	3.90
ASTORE	71.43	33.57	510	01-Jan-1961	31-Dec-2013	2.2	4.4	2.56
KOTLI	73.90	33.52	613	01-Jan-1961	31-Dec-2013	0.3	3.0	13.02
LAHORE	74.40	31.52	215	02-Jan-1960	31-Dec-2013	0.8	0.8	0.68
ASTORE	71.55	32.55	210	01-Jul-1961	31-Dec-2013	3.1	4.5	9.38
MULTAN	71.43	30.20	122	01-Jan-1960	31-Dec-2013	1.2	1.8	1.82
MURREE	73.38	33.92	2167	01-Jan-1960	31-Dec-2013	3.5	3.5	3.29
MUZAFFARABAD	73.48	34.37	701	01-Jan-1961	31-Dec-2013	0.1	0.6	0.52
NAWABSHAH	68.37	26.25	37	01-Jan-1961	31-Dec-2013	2.8	2.0	2.69
NOKKUNDI	62.75	28.82	682	01-Jan-1961	31-Dec-2013	1.8	3.5	2.06
PADIDAN	68.13	26.85	46	01-Aug-1961	31-Dec-2013	2.0	4.6	8.15
PANJGUR	64.10	26.97	980	01-May-1961	31-Dec-2013	2.2	2.1	2.58

PARACHINAR	70.08	33.87	1725	01-Jan-1960	31-Dec-2013	4.7	4.3	5.86
PASNI	63.48	25.27	4	01-Jan-1961	31-Dec-2013	8.8	12.7	22.96
PESHAWAR	71.58	34.02	359	01-Jan-1960	31-Dec-2013	2.6	0.5	0.99
QUETTA	66.88	30.25	1600	01-Jan-1961	31-Dec-2013	9.0	4.0	4.81
ROHRI	68.90	27.70	66	01-Jan-1971	31-Dec-2013	0.7	1.5	0.95
SAIDU SHARIF	72.35	34.73	961	01-Mar-1974	31-Dec-2013	2.3	3.4	3.85
SARGODHA	72.67	32.05	187	01-Jan-1961	31-Dec-2013	0.6	0.7	1.67
SIALKOT	74.53	32.50	251	01-Jan-1960	31-Dec-2013	7.1	10.1	14.58
SIBBI	67.88	29.55	133	01-Jan-1961	31-Dec-2013	4.9	8.8	24.65
SKARDU	75.68	35.30	2209	01-Jan-1961	31-Dec-2013	3.7	1.8	1.41
ZHOB	69.47	31.35	1405	01-Jan-1961	31-Dec-2013	2.0	5.6	6.54
MOENJODARO	68.13	27.32	32	01-Jan-1980	31-Dec-2006	9.3	4.9	17.07
CHITRAL	71.79	35.85	1499	01-Jan-1981	31-Dec-2013	2.6	0.1	0.13
LASBELLA	67.03	24.88	25	01-Jan-1982	31-Dec-2013	2.6	1.4	3.74
LARKANA	68.22	27.57	55	01-Jan-1987	31-Dec-2013	9.9	1.9	2.15
Mangla	73.65	33.14	283	01-Jan-1988	31-Dec-2013	3.1	1.0	1.21
Kamra	72.40	33.75	339	01-Jan-1989	31-Dec-2013	5.4	4.4	4.27
BANNU	70.60	32.99	382	01-Jan-1991	31-Dec-2013	5.8	3.6	2.63
DHAMIAL	73.05	33.61	497	01-Jan-2000	31-Dec-2013	2.6	0.9	0.87
Sukkur	68.84	27.71	68	01-Jan-2000	31-Dec-2013	1.2	2.1	7.05
MandiBahauddin	73.50	32.59	222	01-Jan-2003	31-Dec-2013	0.3	0.0	0.07
D.G.KHAN	70.64	30.05	123	01-Jan-2003	31-Dec-2013	0.3	0.0	0.05
KALAM	72.46	35.48	663	01-Jan-2003	31-Dec-2013	10.0	10.1	10.05
Rawalakot	73.75	33.85	1629	01-Jan-2003	31-Dec-2013	0.2	0.2	0.20
RahimYarKhan	70.30	28.42	80	01-Jan-2003	31-Dec-2013	0.3	0.8	1.23
MalamJabba	72.57	34.80	2267	01-Jan-2003	30-Dec-2013	37.6	38.0	37.62
MirPurKhas	69.01	25.51	21	01-Apr-2004	31-Dec-2013	0.6	0.0	0.06
JHANG	72.32	31.27	156	01-Jan-2008	31-Dec-2013	0.0	0.0	0.05

APPENDIX C: AN INVESTIGATION AND ASSESSMENT OF METEOROLOGICAL AND HYDROLOGICAL DROUGHTS IN PAKISTAN USING HIGH RESOLUTION IN SITU GRIDDED DATA SET

(To be Submitted to Geophysical Research Letters)

Furrukh Bashir

Department of Hydrology and Atmospheric Sciences, University of Arizona, Tucson, AZ, USA

Pakistan Meteorological Department, Islamabad, Pakistan

Xubin Zeng, Hoshin Gupta, Pieter Hazenberg, Thomas Galarneau

Department of Hydrology and Atmospheric Sciences, University of Arizona, Tucson, AZ, USA

ABSTRACT

Agriculture based economy of Pakistan is vulnerable to drought, that occurs frequently across the country on seasonal and annual scales. Long term droughts that may span for 4-5 years occur nearly every 17 years in Pakistan. We have reconstructed history of drought in Pakistan using in situ observations based high resolution gridded data through SPI methodology on different time scales. Furthermore, we have explained the transition of meteorological drought to hydrological drought using river inflows data of large rivers of Pakistan, and explained the sensitivity of rivers' inflows in Pakistan to rainfall and temperature of different seasons. On the basis of this analysis, we have proposed a solution i.e., construction of water reservoirs to tap water resources from northern mountains that are less sensitive to dominant rainfall variability in Pakistan, so that it may buffer the impact of droughts in southern parts of Pakistan. Furthermore, we have demonstrated the potential of PDSI as an operational tool for drought monitoring in Pakistan.

Keywords: Drought, SPI, SRI, PDSI, Pakistan.

1. INTRODUCTION

Water is integral to ability of an economy to produce goods and deliver services. In Pakistan agriculture is the main driving force of economic activities as it provides basic ingredients for sustenance and raw material to industry. Increase in population has created a greater need of growing food that put more pressure on the water resources by increasing its demand. With changing climate and global warming, the agriculture is becoming more vulnerable to climatic hazards, such as, droughts. Though, drought is a recurring feature of climate (Bordi and Sutera, 2007), as it occurs in all climatic zones (Mishra and Singh, 2010), regardless of how much precipitation an area receives climatologically and it is different from aridity that is confined to climatologically low rainfall regions. However, it is related to the reduction in the amount of precipitation received over an extended period, that may range from months to years. The hydrometeorological variables that may play a role in occurrence of drought includes,

temperature; high-winds; low relative humidity; low cloud cover; timing and characteristics of rain, including frequency of rainy days during crop growing seasons, intensity and duration of rain, along with onset and termination of rainy season (such as monsoon). Heat wave is sometimes confused with drought, however, the typical time scale of the former is on the order of week and later may persist for years (Chang and Wallace, 1987). Though, the combination of both has dire socio-economic consequences (Mishra and Singh, 2010). Drought affects both surface and groundwater resources and can lead to decrease in water supply, deteriorate water quality, fail agriculture and farming, decrease hydro-power generation, disturb riparian ecology and may halt recreational activities. Moreover, it severely affects transport of nutrients and organic matter associated with movement of water through several pathways.

Among all the meteorological hazards, droughts rank the first, and they differ from other in several ways; i) droughts are possibly the most slowly developing hazards, ii) that may last for a very long duration, iii) they are least predictable with accuracy, and iv) their far-reaching impacts may affect the economy beyond the area-experiencing the drought. For instance, recent crisis of Syria is ascribed to severe multiyear drought beginning in the mid-2000s combined with inefficient and often unmodernized irrigation systems and water abstractions by other parties in the eastern Mediterranean, including especially Syria, contributed to the displacement of large populations from rural to urban centers, food insecurity for more than a million people, and increased unemployment with subsequent effects on political stability (Gleick, 2014). It followed by an unprecedented large and abrupt flow of war refugees from the Middle East to Europe and other continents (Heisbourg, 2015). Therefore, unlike floods, hurricanes, earthquakes and tornedoes a drought affects water resources and seldom results in structural damages. For the same reason the quantification of its impact and the provision of relief are far more difficult for drought than for other natural hazards (Wilhite, 2000).

Bryant (2005) ranked hazard events considering their characteristics and impacts. Key hazard characteristics used for ranking included the degree of severity, the length of event, total areal extent, total loss of life, total economic loss, social effect, long-term impact, suddenness,

and occurrence of associated hazards. It was found that drought stood first based on most of the hazard characteristics.

Inter-governmental Panel on Climate Change (IPCC) study suggests that production of rice, maize, and wheat in the past few decades has declined in many parts of Asia due to increase in water stress, arising partly from increasing temperature, increasing frequency of El Nino events and reduction in the number of rainy days (Bates *et al.*, 2008). For instance, during the drought of 1999-2000, up to 60 million people in central and southwest Asia were affected by a persistent multi-year drought, that was considered as the largest one from a global perspective, in which Iran, Afghanistan, Western Pakistan, Tajikistan, Uzbekistan and Turkmenistan experienced the severe impacts.

(Xie *et al.*, 2013), using standardized precipitation Index (SPI) has analyzed the spatiotemporal variability of drought in Pakistan for the period of 1960-2007 using APHRODITE (Yatagai *et al.*, 2012) data set. They found that droughts are wide-spread and often occur simultaneously over large areas. They identified three intensive drought episodes; i) the late 1960s to early 1970s; ii) the middle 1980s, and iii) the late 1990s to early 2000s and they identified a 16-year drought recurrence period. Another study (Adnan *et al.*, 2016) points out a drought episode in 1952, in addition to the above mentioned ones, and found 2000-2002 drought as the most widespread (covering almost whole south central Asia, CSA) and longest in the recorded history. Another study (Haroon *et al.*, 2016) on Pakistan identified same three episodes (1968-1975, mid-1980s and 1999-2003) with better identification of duration using SPI. However, none of them have explicitly estimated the intensity and severity of the drought. Quantification of duration is also vague as it indicates onset and ending in terms of years not in terms of months.

Water resources availability in Pakistan is dominated by the snow and glacier melt from Hindukush, Karakoram and Himalayan (HKH) mountain ranges situated in the north of the country, where due to increase in humidity and cloud cover less glacier melt is available in summer season when its demand is highest (Bashir *et al.*, 2017). Therefore, for an agriculture economy like Pakistan, an objective evaluation of drought condition is the first step toward

planning of water resources management to prevent and mitigate the negative impacts of future occurrences. For this purpose, several indexes have been developed to evaluate the water supply deficit in relation to the time duration of precipitation shortage (Bordi and Sutera, 2007; Mishra and Singh, 2010). Among them we select Standardized Precipitation Index (SPI), the Palmer Drought Severity Index (PDSI), and Standardized River Index (SRI) to evaluate the onset, duration, intensity, and severity in Pakistan from 1960-2013 using very high resolution in situ observed gridded data PAK-HYM-1.0 (Bashir et al, 2018, in review). Moreover, we use SPI methodology to identify the areal coverage and spatial spread of drought in Pakistan.

2. DATA AND STUDY AREA

Pakistan covers southwestern portion of South Asia. It extends latitudinally from 23.45ºN to 36.75ºN, and longitudinally from 61ºE to 75.5ºE. Pakistan is surrounded by Hindukush, Karakoram and Himalayan mountain ranges in the north. In the northeast, it has a common border (595 km) with China. India is in the east of Pakistan with 909 km long boundary. Arabian sea is situated in the south of Pakistan and Iran is in the west. Afghanistan is in the northwest of Pakistan and shares 2240 km long common boundary. Nearly 55% of Pakistan is covered by mountains and plateaus, whereas, 45% is covered by plains and deserts (see Figure C.1a for hypsometry of Pakistan)

Climatologically, Pakistan can be divided into three regions: i) areas that receive precipitation mainly in winter season, such as, north and northwestern mountainous locations; ii) areas that receive precipitation mainly in summer due to monsoon such as south and southeast, and iii) areas that receive varying amount of precipitation in both seasons such as River Indus plains in central Pakistan (annual mean daily precipitation, annual mean daily maximum and minimum temperature are presented in Figures C.1b, C.1c and C.1d, respectively). Spring and autumn seasons are generally dry throughout the country. Spring season is warm in plains and summer is warm in mountainous areas. Furthermore, dry locations that are away from the Arabian Sea are the warmest.

2.1 HIGH RESOLUTION GRIDDED HYDROMETEOROLOGY DATA FOR PAKISTAN

The reliability of drought analyses carried out by applying any drought indexes strongly depends on the quality of the primary data. Specifically, in assessing dry (wet) periods, it is highly desirable to have at hand a data set that: (i) is easy to access; (ii) uniformly covers the whole region; (iii) has better spatial resolution; (iv) has a time-duration sufficiently long to be trustworthy in a statistical sense and (v) is optimal in the sense of capturing consistently dry and wet events.

Therefore, meteorological studies favor “analysed data”, i.e. a set of observations which have been processed through several quality checks and objective analyses. The final products of this procedure are uniformly gridded fields on a regional or global scale of hydrometeorological variables that are released for further applications. Here, we use PAK-HYM-1.0, which is a daily 0.05°x0.05° gridded maximum and minimum temperature, and precipitation data set for Pakistan based on 67 sites for the period of 1960-2013 (Bashir, 2018). This dataset uses two to four times more meteorological stations than any other similar product, with continuous time series of observations. Availability of both temperature and precipitation at same grid-resolution enables better quantification of drought.

2.2 RIVER INFLOWS DATA

Inflows data of different rivers of Pakistan is analyzed to relate meteorological drought with river inflows by using Standardized Runoff Index (explained below). The name of the rivers, location of their gauging stations (see Figure C.2) and duration of data is presented in Table C.1.

3. METHODS

3.1 STANDARDIZED PRECIPITATION INDEX (SPI)

Usually drought monitoring indexes are based on precipitation amount, and they measure the deviation of actual precipitation from a historically established norm. Some of them, instead, utilize other climatological variables such as temperature, evapotranspiration or soil moisture, into account, as well. However, if we wish to compare drought conditions of different areas, which often have different hydrological balances, the most important characteristic of an index is its standardization. For this purpose, the SPI seems to be the most powerful index. It is a standardized index, and can be computed on different time scales, so as to allow to monitor all drought conditions ranging from meteorological, agricultural, hydrological and socioeconomic or even ground water drought.

In SPI methodology, we apply a distribution (such as gamma or extreme value distribution) on precipitation (accumulated precipitation in case of 3-, 6-, 12-, 24-months) time series. The cumulative probability is transformed to the standard normal random variable Z with mean zero and variance of one. This is an equiprobability transformation in which a variate from one distribution (such as gamma) to a variate with a distribution of prescribed form (i.e., normal distribution) such that the probability of being less than a given value of the variate shall be the same as the probability of being less than the corresponding value of the transformed variate (Mckee *et al.*, 1993). This method is illustrated in Figures C. 3a and C. 3b, where a 3-month precipitation amount (**O**ctober-**N**ovember-**D**ecember) of Islamabad in millimeters is converted to SPI value with a mean of zero and variance of one. Figure 3a presents cumulative probability of **OND** Islamabad precipitation for the period of 1960-2013 (Blue line) and its magnitude is presented on x-axis in mm and empirical cumulative probability of events of various magnitudes are presented on the y-axis. Black Line in Figure C. 3a indicates the cumulative probability of the fitted gamma distribution. Transformation to empirical cumulative probability in SPI framework is presented in Figure C. 3b where blue line indicates observations and black line presents

theoretical fit of normal cumulative probability distribution. Please see S.I. for the details of SPI-methodology.

3-, 6-, 12-, and 24-month SPI are calculated at each grid box within in boundaries of Pakistan using PAK-HYM-1.0 precipitation data. Figures C. 4a1, C. 4a2, C. 4a3 and C. 4a4 presents the dominant mode of variability of precipitation in Pakistan computed from 3-, 6-, 12-, and 24-month SPI, respectively.

3.2 KOLMOGOROV-SMIRNOV GOODNESS OF FIT TEST

We use Kolmogorov-Smirnov Goodness of Fit Test, to determine if observations follow the Gamma distribution (see Supporting Information). It is a non-parametric hypothesis test that does not assume any particular underlying distribution (Massey, 1951). We use two sample-test with following hypothesis:

H_0: the data comes from the Gamma distribution.

The two-sided test uses the maximum absolute difference between the cumulative distribution functions of the distribution of observations and Gamma distribution. The test statistic is:

$$D^* = \max\left(\left|\hat{F}_1(x) - \hat{F}_2(x)\right|\right) \quad (1)$$

In two-sided Kolmogorov-Smirnov goodness of fit test, the decision to reject the null hypothesis is based on comparing the *p*-value with the significance level i.e., 0.05 and 0.01 in this case.

Figure C. 3c presents the percentage of rejection rates of the Kolmogorov-Smirnov goodness of fit test for gridded precipitation of PAK-HYM-1.0 data. We have applied the test on 1-, 3-, 6-, 12-, and 24-months accumulated precipitation on monthly scale. As expected highest rejection came from 1-month time series, whereas lowest came from 12-month accumulated precipitation. This test suggests that gamma distribution is good fit for the gridded precipitation data provided by PAK-HYM-1.0. please note that K-S goodness of fit test rejects values above a

significance level, therefore, it rejects less values at 0.01 in comparison to values rejected at 0.05 significance level.

3.3 MAXIMUM COVARIANCE ANALYSIS

Empirical Orthogonal Function (EOF) seek structures in the data that explains the maximum amount of variance in two-dimensional data sets. In one dimension of the data we are seeking structure and in the other dimension the realization of this structure are sampled. To seek spatial structures that vary with time, for instance, space will be used as the structure dimension and time as sampling dimension. Resulting analysis will provide a set of structures in the spatial dimension along with a complimentary set of structure in sampling dimension (i.e., time). Structure in time dimension are called Principal Components (PCs). We have utilized MATLAB built-in function called 'SVD' (https://www.mathworks.com/help/matlab/ref/svd.html) to computer EOF and PC of the gridded SPI timeseries from PAK-HYM-1.0.

The statistical significance of a PC (time domain) is estimated by the degree of separation between its eigenvalues. Closer eigenvalues indicates a large inter-sample variability, whereas, well separated eigenvalues indicates a substantial amount of variability among the gridded data is explained by EOF or PC related to that specific eigenvalue (North *et al.*, 1982). The degree of separation between an eigenvalue and its adjacent neighbor in the diagonalized covariance matrix is given by following expression:

$$\Delta\lambda_i = \lambda_i \sqrt{\frac{2}{N}} \qquad (2)$$

where $\Delta\lambda_i$ is the error bar on eigenvalue and N is the effective number of degrees of freedom in the data that we computed using autocorrelation in the timeseries. The eigenvalues and corresponding pair of EOF/PC are considered significant if the error bars (Figure C. 4c1, C. 4c2, C. 4c3 and C. 4c4) do not overlap with others. If first eigenvalue is significantly different from the successive ones, then it is considered to be explaining most of the variance and can be used for the reproducibility of the data and it indicates its robustness, as well. As first EOF is

equivalent to the projection of the data matrix onto the first PC, therefore, correlation of first PC to the original SPI at each grid point will deliver same structure as of first EOF. However, the resulting map will indicate the correlation between the different gird boxes and PC representing maximum variability (see Figure C. 3b1, C. 3b2, C. 3b3, C. 3b4). Higher order EOF/PC will not reflect important information as they do not present sufficient variance in them.

3.3.1 PERIODOGRAM

A time series can be expressed as a combination of cosine or sine waves with different periods and amplitudes, that indicates the periodic behavior of the phenomena. The important frequencies (periods) in the times series can be filtered through periodogram. The periodogram indicates a measure of the relative importance of possible frequency values that might explain the oscillation pattern of the observed data.

Power Spectral Density (PSD) of periodogram describes the power present (distribution of variance of series) in the signal as function of frequency, per unit frequency. PSD is of interest because it can be helpful to compare the magnitude of variance of phenomena occurring at same frequency. We have computed the power spectrum of 3-, 6-, 12- and 24-month SPI in Figure C. 4d1, C. 4d2, C. 4d3 and C. 4d4, respectively. Moreover, we have computed same for 12-month SRI (described below) in Figure C. 5 for different rivers. We have utilized MATLAB function (https://www.mathworks.com/help/signal/ref/periodogram.html?searchHighlight=periodogram&s_tid=doc_srchtitle) for periodogram analysis (Trauth, 2007: 21).

3.4 STANDARDIZED RUNOFF INDEX (SRI)

This index is based on the concept of Standardized Precipitation Index (SPI), as discussed earlier. Shukla and Wood (2008) derived Standardized Runoff Index (SRI) which incorporates hydrologic processes that determine the seasonal loss in streamflow due to the influence of climate. As a result, month to seasonal time scales SRI is a useful complement to SPI for identification of hydrological aspects of droughts (Mishra and Singh, 2010). The methodology to

compute SRI is same as of SPI, expect it utilizes streamflow data instead of precipitation data. The 3-, 6-, 12- and 24-month SRI of river Astore, Hunza, Kabul, Indus, Jhelum, Chenab, Ravi and Sutlaj are presented in Figure C. 5.

3.5 PALMER DROUGHT SENSITIVITY INDEX (PDSI)

The PDSI is based upon a set of empirical relationships derived by (Palmer, 1965) to express regional moisture supply standardized in relation to local climatological norms. It has been used widely in the United States of America since its introduction in 1965. The index is a sum of the current moisture anomaly and a fraction of the previous index value. The moisture anomaly is defined as

$$d = P - \hat{P} \tag{3}$$

where P is the total monthly precipitation, and $\hat{P}$ is the precipitation value 'climatologically appropriate for existing conditions' (Palmer 1965). $\hat{P}$ represents the water balance equation defined as

$$\hat{P} = \overline{ET} + \overline{R} + \overline{RO} - \overline{L} \tag{4}$$

where $\overline{ET}$ is the evapotranspiration, $\overline{R}$ is the soil water recharge, $\overline{RO}$ is the runoff, and $\overline{L}$ is the water loss from the soil. The overbars signify that these are average values for the given month taken over some calibration period. $\hat{P}$ is a hydrological factor and needs to be parameterized locally. The four potential values $\overline{PE}, \overline{PR}, \overline{PRO}$ and $\overline{PL}$ are used to compute four coefficients which are dependent on the climate of areas under consideration:

$$\alpha_j = \frac{\overline{ET}_j}{\overline{PE}_j},$$

$$\beta_j = \frac{\overline{R}_j}{\overline{PR}_j},$$

$$\gamma_j = \frac{\overline{RO}_j}{\overline{PRO}_j},$$

$$\delta_j = \frac{\overline{L}_j}{\overline{PL}_j},$$

Where $j = 1, \ldots, 12$, and overbar indicates that the coefficients are computed using average values of the month j. Therefore,

$$\hat{P} = \alpha_j \overline{PE} + \beta_j \overline{PR} + \gamma_j \overline{PRO} - \delta_j \overline{PL}$$

Here $\hat{P}$ is analogous to simple water balance where precipitation is equal to evapotranspiration plus runoff and ground water recharge and plus or minus any change in soil moisture storage.

The Palmer moisture anomaly index (Z index) is then defined as

$$Z = Kd \tag{5}$$

and the PDSI for month i is defined as

$$PSDI_i = 0.897\, PDSI_{i-1} + \frac{Z_i}{3} \tag{6}$$

K acts as a climate weighting factor and is applied to yield indices with comparable local significance in space and time. The resultant PDSI values are broken down into 11 categories, ranging from extremely dry to extremely wet. These are listed in Table S2. As implied in the above description, the PDSI is usually calculated over a monthly period. However, there is nothing to prevent calculations across other time periods, e.g. weekly or bi-monthly. PDSI methodology is generally not clear and is hard to maintain consistency with the procedure described by Alley (1984) and Palmer (1965), and various codes available online for the calculation of the PDSI lacks transparency. Therefore, we are using a MATLAB code provided by Jacobi *et al.*, (2013) for the consistency and reproducibility.

4. RESULTS AND DISCUSSION

Kolmogorov-Smirnov (K-S) goodness indicates that Gamma distribution is a good fit for the PAK-HYM-1.0 precipitation data. Almost 6% grid points are rejected by K-S test on 1-month scale on significance level of 0.05, therefore, we are not using 1-month SPI analysis for the identification of droughts in Pakistan. However, 3-months, 6-months, 12-months and 24-months accumulated precipitation are performing better, as less number of grids points are reject in them. The grid points that are reject by the K-S test are situated in extremely dry locations where zero values are dominating the time series and an occasional single value with high magnitude exists. The main reason that 24-months accumulated precipitation has higher rejection percentage than 12-month is the same as consecutive two wet years in an extremely dry area will make gamma distribution unfit, even though the same two years on annual scale are suitable for distribution fitting. Overall, data are suitable for gamma distribution fit that is further transformed into normal distribution using equiprobability distribution as presented in Figures C. 3a and C. 3b, where cumulated probability distribution of October-November-December (OND) accumulated precipitation of Islamabad is overlaid by the cumulative gamma distribution, and its equiprobability normal distribution is presented in next panel, where it is overlaid with cumulative normal distribution that presents a good fit.

Employing maximum covariance analysis (described earlier) we computed Principal Components (PC) on 3-month, 6-month, 12-month and 24-months SPI time series of each grid point (See Figures C. 4a1, C. 4a2, C. 4a3, and C. 4a4), furthermore, their eigenvalue spectrum with significance is presented in (Figures C. 4c1, C. 4c2, C. 4c3, and C. 4c4) for each SPI, respectively. Eigenvalue spectrum indicates that PC-1 is significantly separated in all SPI time series expect in 24-month SPI, therefore, we use PC-1 of all SPI to present the wet and dry periods in Pakistan. PC-1 time series for 3-month SPI indicates recurring episodes of dry spells from 1960-1970 with varying intensity, when most of the time the magnitude remained below -1 (mild drought to severe drought). 1960, 1962, 1963 1965, 1967-69, 1971, 1973, 1977, 1985, 1991, 2000-2002, 2004, 2006 are the years when severe to extreme drought are experienced in Pakistan. 39 incidences of severe to extreme droughts events are recorded in 3-month SPI

analysis of 54-year data that indicates that 6% of the times Pakistan is experiencing severe to extreme meteorological drought in this period.

Winter season drought is prominent in 1960-1973, 1984-1989, 1999-2002, 2006 and 2010-2013. Spring season drought is prominent in 1968-1975, 1988, 1999-2005, 2010. Summer season drought is prominent in 1960-1974, 1991, 1999-2000, and 2004. Whereas, autumn season is dominated from 1960-1974, 1985-1986, 1999-2002.

Winter and spring season witnessed severe drought in 1960-1974, 1984-1989 and 1999-2003, whereas, summer and autumn experienced mild drought in same years. On the other hand, 1991 was wet year in winter and spring and dry year in summer and autumn. In nutshell, 3-month SPI indicates that all seasons have 3 severe drought ($< -1.49\ SPI$) episodes except autumn that has 1. Whereas 6-month SPI shows winter has 2, spring has 5, summer has 2 and autumn has 3 drought episodes in 54 years.

1984-1989 and 1999-2004 droughts were stronger and longer in winter and autumn in comparison to summer and autumn. On the other hand, 1960-1975 drought episode is more prominent in spring and autumn then winter and spring.

Correlation map of PC-1 and original SPI time series is presented in Figure C. 4b1 that indicate that drought is mainly prevailing in southern Punjab, most of the Sindh province and eastern Baluchistan. As stated earlier these are the areas that are influenced mainly by the monsoon rainfall. For instance, mountainous Gilgit-Baltistan (north) and western hilly locations of Baluchistan are not very much influenced by the dominant variability of precipitation in Pakistan as these areas receive a portion of their annual precipitation from the western disturbances, as well. Therefore, it can safely assume that meteorological droughts (seasonal) in Pakistan are dominated by the variability in monsoon rainfall in Pakistan, however, western disturbances (attributed to Rossby wave trains) are also important in long term droughts. Spatial coverage of the drought in Pakistan suggests that variability is dominated by the influence of the areas that receive rainfall in monsoon season predominantly (see Figure C. 4b1, C. 4b2, C. 4b3 and C. 4b4).

Periodogram analysis of the 3-months SPI (Figure C. 4b1) indicates that most of the variability (power) is present at 85.3-year cycle (5 times of 17.1 that is second prominent frequency). However, 12- and 24-month SPI periodogram point 17-year cycle as the strongest one and it has more power than same frequency in 3 - and 6- months periodograms. It indicates that at annual scale Pakistan is vulnerable to a recurrence of severe drought every 17 years. Right now, in year 2017, we are almost at the brink of another extreme drought period that may extend for 4 to 5 years. Moreover, on annual scales winter droughts are contributing more than summer droughts.

For the assessment of hydrological droughts, we have examined the inflow data of 8 rivers using SRI (discussed earlier) and results are presented in Figure C. 5. Astore (Figures C. 5a1, C. 5b1, C. 5c1, and C. 5d1) runoff is in a mountainous catchment dominated by snow and glacier melt along with precipitation. It indicates dry period in 1970s and 1980, and around 2000, whereas 1990s was a wet period. 3-month SRI shows a very mild dry spell in river runoff in 1994 that is discernable in SPI analysis. Therefore, Astore catchment is mainly regulated by the rainfall. In contrast Hunza runoff indicates drought in 1990 when most of the Pakistan experienced wet spell. It means that runoff in Hunza catchment is more sensitive to the seasonal temperature then same season's precipitation, however, it is dependent on the previous seasons precipitation present in the form of snow and ice. Therefore, when there is more precipitation, cloudiness will cease melting and less runoff will be available in rive Hunza. For similar reason in 1960s and 1970s when most of the country was experiencing dry conditions river Hunza delivered runoff from melting snow and ice. Winter of 1998 was meteorologically very wet and it is marked as dry in Hunza river.

River Kabul, originating from northwestern part of Pakistan, enters Afghanistan where it joins several other tributaries and then flow eastward and enters Pakistan to join River Indus. Its runoff is composed of snow and glacier melt and rainfall. SRI analyses (Figures C. 5a3, C. 5b3, C. 5c3, C. 5d3 and C. 5e3) of river Kabul indicate that its runoff is sensitive to precipitation in Pakistan as it shows wet spell in early 1990s (dominated by flooding in Pakistan) and dry spell

after 2000 (dominate by drought in Pakistan). Periodogram analysis indicates high variance at 21 and 14-year cycles.

River Indus inflows at Tarbela (Figures C. 5a4, C. 5b4, C. 5c4, C. 5d4 and C. 5e4) is also comprised of snow and glacier melt along with rainfall. Overall response of the catchment is more sensitive to autumn and winter season rainfall variability and there is a lag between meteorological drought and its transition to hydrological drought in River Indus at Tarbela. For instance, the meteorological drought that started in May 1999 is reflected for the first time in Indus inflows in Apr 2000. The most dominant variation (drought) recur in Indus river every 21 years according to the available data set.

River Jhelum inflows at Mangla (Figure C. 5a5, C. 5b5, C. 5c5, C. 5d5 and C. 5e5) is more sensitive to summer and autumn rainfall than any other season. 12-month SRI indicates that drought started in May 1999 and terminated in Jun 2003 that is similar to 12-month SPI, most of the variance recur at 14 years.

Among three western rivers whose water is allocated to Pakistan under Indus Water Treaty of 1960, Chenab's catchment is extended in farthest east. For same reason, it is sensitive to all season as it receives snow in winter and early spring, and monsoon in summer in its lower reaches. Interestingly, its response to 2000s drought is relatively milder than other rivers (Figure C. 5a6, C. 5b6, C. 5c6, C. 5d6 and C. 5e6).

According to Indus Water Treaty the two eastern rivers of Ravi and Sutlaj are allocated to India for abstractions. Since then waterflow to Pakistan in these rivers is continuously decreasing. Nevertheless, when it is drought, they will also indicate it (as presented in lowest two panels of each column of Figure C. 5).

Figure C. 6 presents PDSI for different sites in Pakistan. Most of the soil in Pakistan is sandy loam. We have assumed an available water capacity of almost 660 mm/m., for each station (this number will be updated after obtaining a digital soil map from survey of Pakistan). PDSI is estimated for each site on monthly basis. Almost all locations are indicating a varying severity of drought around year 2000. As PDSI is water budget estimation and it is highly sensitive to soil moisture, therefore, it is more viable for operational drought monitoring instead of historical

drought identifications where its capability is compromised due to limited availability of historical soil moisture.

The operational implementation of PDSI is presented in Figure C.7, where a contour map of PDSI for Jul 2001 for whole Pakistan is presented. It is clear from the Figure C.7 that PDSI can be adopted on fortnight to monthly scales to identify agricultural drought in different locations of Pakistan with diverse climatology.

5. CONCLUSIONS

With increase in population and global warming, pressure on the water resources is increasing, and water consuming sectors such as agriculture and hydropower generation are becoming more vulnerable to climatic hazards, such as, drought, that may last for a long time. It is least predictable with limited structural damage, and has far-reaching consequences on the socio-economic life of people beyond its own physical geographic extent. It is ranked first among natural hazards. Droughts in Pakistan are wide-spread and often occur simultaneously on large areas, sometimes, throughout the country. Almost four long term drought episodes are identified in the history of Pakistan, however, due to the nature of the data and nature of the drought indexes it is hard to evaluate their severity in absolute terms, and most of the past studies on Pakistan mainly discussed the meteorological drought on different timescales. Therefore, to bridge this gap we present analyses of both meteorological drought and hydrological drought in Pakistan.

We have used high resolution precipitation and temperature data PAK-HYM-1.0 that covers the period of 1960-2013 on 0.05º latitude-longitude (Bashir, 2017).

We have utilized SPI to identify the meteorological drought with its spatio-temporal variations in Pakistan. We have applied Kolmogorov-Smirnov goodness of fit test to evaluate the suitability of gamma-distribution for accumulated precipitation from PAK-HYM-1.0 and found that it is a good fit for the precipitation data.

Maximum covariance analysis on 3-, 6-, 12-, and 24-months gridded SPI time series across the Pakistan provides PC-1 that explains maximum variability. PC-1 of 3-months SPI indicates that Pakistan remained in severe to extreme drought for more than 6% of its 54 years

long time series . Seasonal droughts are dominated in summer and autumn seasons, whereas, long-term droughts are accompanied by winter and spring seasons. 17 years/cycle is the dominant frequency for the occurrence of drought in Pakistan, and areas of southern Punjab, eastern Baluchistan and Sindh are most vulnerable to drought. Considering, 17 years/cycle for drought, a long-term and intense drought in Pakistan that may persist for 4 to 5 years is expected in the next few years.

SRI of river inflows indicates that rivers originating from HKH mountains are sensitive to winter and spring season precipitation, and summer and autumn season temperatures. High elevated catchments such as Hunza is more sensitive to temperature then precipitation, however, other relatively low-lying catchment are sensitive to winter season precipitation and areas further low in elevation are vulnerable to summer and autumn season precipitation as well. High altitude catchments such as Hunza are good buffer against dry periods, as they provide water in meteorologically dry season and holds water in meteorologically wet seasons. Western most catchment of Kabul covers a vast area and therefore sensitive to precipitation and temperature according to its hypsometry, however, its response at gauging station in Pakistan is more sensitive to precipitation in Pakistan. In addition to this, a gradual decrease in inflows of Kabul river is witnessed in time series that indicates potential human influence in upstream areas (Darunta Dam since 1960, and Naghlu Dam since 1968, Source: https://en.wikipedia.org/wiki/List_of_dams_and_reservoirs_in_Afghanistan).

River Indus presents integrated response of Upper Indus Basin (UIB) comprised of HKH mountains in Pakistan. It is more sensitive to autumn and winter season rainfall and serves as a buffer in transition of meteorological drought to hydrological drought, similar to Hunza catchment, but up to a lesser extent as its low-lying areas are sensitive to spring and summer season precipitation as well. Meteorological drought that was started in May 1999 was first reflected after a year in River Indus at Tarbela in April 2000. River Kabul and River Indus both have drought recurring cycle of almost 21 years.

River Jhelum at Mangla is more sensitive to summer and autumn rainfall. Its buffer effect is lesser then River Indus at Tarbela. River Chenab catchment is extended in to lesser Himalayas

in east and therefore its meteorological forcing are relatively different from Pakistan, for same reason, its response to the drought episode around year 2000 is relatively milder than other rivers.

The other eastern rivers such as Ravi and Sutlaj are regulated by India since Indus Water Treaty, therefore, their runoff is decreasing since 1960. However, the year 2000 drought episode is prominent in these rivers.

Since, PDSI is very sensitive to available water capacity (AWC) as it is based on soil water budget, therefore, we have assumed a constant AWC for all sites in Pakistan and have presented its strength as a real time operational tool. Since PAK-HYM-1.0 is a daily product that can be upgraded on daily basis, as well. Therefore, it can be adopted to estimate the drought conditions throughout Pakistan using PDSI on desirable timescale ranging from a week to a month.

Since Pakistan is vulnerable to drought on seasonal to inter-annual scales, and meteorological drought is translated to hydrological drought very quick in its southern areas. Pakistan needs to consider development of more structure that can tap water from HKH mountains. River water from HKH mountains is acting as a buffer against meteorological drought and water storage on River Indus can help Pakistan to ameliorate drought in its southern areas. Small dams are also recommended to tap rainfall water on seasonal scale that can be used for irrigation, however, small dams cannot act as buffer against inter-annual droughts.

ACKNOWLEDGMENTS, SAMPLES, AND DATA

We thank Ghulam Rasul, Director General, and other Directors, Pakistan Meteorological Department (PMD) (http://www.pmd.gov.pk/) for providing Precipitation, Temperature and River Inflows data. Furthermore, we are thankful to United States Education Commission Pakistan (USEFP) for providing Fulbright scholarship to Furrukh Bashir to conduct this research. Moreover, we acknowledge, all the datasets utilized in this study by citing them.

REFERENCES

Adnan S, Ullah K, Shouting G. 2016. Investigations into precipitation and drought climatologies in south central Asia with special focus on Pakistan over the period 1951-2010. Journal of Climate 29(16): 6019–6035. DOI: 10.1175/JCLI-D-15-0735.1.

Alley WM. 1984. The Palmer Drought Severity Index: Limitations and Assumptions. Journal of Climate and Applied Meteorology, 1100–1109. DOI: 10.1175/1520-0450(1984)023<1100:TPDSIL>2.0.CO;2.

Bashir F, Zeng X, Gupta H, Hazenberg P. 2017. A Hydro-meteorological Perspective on the Karakoram Anomaly using Unique Valley-based Synoptic Weather Observations. Geophysical Research Letters. DOI: 10.1002/2017GL075284.

Bates BC, Kundzewicz ZW, Wu S, Palutikof JP. 2008. 2008: Climate Change and Water. Technical Paper of the Intergovernmental Panel on Climate Change, IPCC Secretariat. Climate change and water. DOI: 10.1016/j.jmb.2010.08.039.

Bordi I, Sutera A. 2007. Drought Monitoring and Forecasting at Large Scale. Methods and Tools for Drought Analysis and Management, 3–27. DOI: 10.1007/978-1-4020-5924-7_1.

Bryant E. 2005. Natural Hazards. Cambridge University Press. DOI: 10.1017/CBO9781107415324.004.

Chang F-C, Wallace JM. 1987. Meteorological Conditions during Heat Waves and Droughts in the United States Great Plains. Monthly Weather Review, 1253–1269. DOI: 10.1175/1520-0493(1987)115<1253:MCDHWA>2.0.CO;2.

Gleick PH. 2014. Water, Drought, Climate Change, and Conflict in Syria. Weather, Climate, and Society 6(3): 331–340. DOI: 10.1175/WCAS-D-13-00059.1.

Haroon MA, Zhang J, Yao F. 2016. Drought monitoring and performance evaluation of MODIS-based drought severity index (DSI) over Pakistan. Natural Hazards. Springer Netherlands 84(2): 1349–1366. DOI: 10.1007/s11069-016-2490-y.

Heisbourg F. 2015. The Strategic Implications of the Syrian Refugee Crisis. Survival 57(6): 7–20. DOI: 10.1080/00396338.2015.1116144.

Jacobi J, Perrone D, Duncan LL, Hornberger G. 2013. A tool for calculating the palmer drought indices. Water Resources Research 49(9): 6086–6089. DOI: 10.1002/wrcr.20342.

Massey FJ. 1951. The Kolmogorov-Smirnov Test for Goodness of Fit. Journal of the American Statistical Association 46(253): 68–78. DOI: 10.1080/01621459.1951.10500769.

Mckee TB, Doesken NJ, Kleist J. 1993. The relationship of drought frequency and duration to time scales. AMS 8th Conference on Applied Climatology (January): 179–184. DOI: citeulike-article-id:10490403.

Mishra AK, Singh VP. 2010. A review of drought concepts. Journal of Hydrology. Elsevier B.V. 391(1–2): 202–216. DOI: 10.1016/j.jhydrol.2010.07.012.

North GR, Bell TL, Cahalan RF. 1982. Sampling Errors in the Estimation of Empirical Orthogonal Funtions. Monthly Weather Review, 699–706. DOI: 10.1175/1520-0493(1982)110<0699:SEITEO>2.0.CO;2.

Palmer WC. 1965. Meteorological Drought. U.S. Weather Bureau, Res. Pap. No. 45, 58.

Trauth MH. 2007. MATLAB® recipes for earth sciences: Second edition. MATLAB® Recipes for Earth Sciences (Second Edition). DOI: 10.1007/978-3-540-72749-1.

Wilhite DA. 2000. Chapter 1 Drought as a Natural Hazard: Concepts and Definitions. Drought Mitigation Center Faculty Publications. Paper 69. 3–18.

Xie H, Ringler C, Zhu T, Waqas A. 2013. Droughts in Pakistan: a spatiotemporal variability analysis using the Standardized Precipitation Index. Water International. Routledge 38(5): 620–631. DOI: 10.1080/02508060.2013.827889.

Yatagai A, Kamiguchi K, Arakawa O, Hamada A, Yasutomi N, Kitoh A. 2012. Aphrodite constructing a long-term daily gridded precipitation dataset for Asia based on a dense network of rain gauges. Bulletin of the American Meteorological Society 93(9): 1401–1415. DOI: 10.1175/BAMS-D-11-00122.1.

Table C. 1. Names, gauging location and duration of data of different rivers used in the study.

River	Gauging Location	Duration
Astore	Doyian	Jan 1974 – Dec 2003
Hunza	Dainyour	Jan 1966 – Dec 2010
Kabul	Nowshehra	Apr 1961 – Mar 2005
Indus	Tarbela	Apr 1961 – Mar 2005
Jhelum	Mangla	Apr 1961 – Mar 2005
Chenab	Marala	Apr 1961 – Mar 2005
Ravi	Balloki	Apr 1961 – Mar 2005
Sutlaj	Qila Ganda Singh	Apr 1961 – Mar 2005

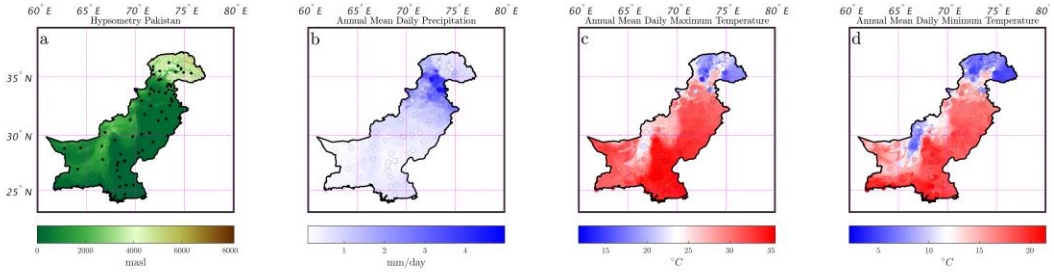


Figure C. 1. Hypsometry of Pakistan with the location of meteorological observatories (a), annual mean daily precipitation (b), annual mean daily maximum temperature (c) and minimum temperature (d) of Pakistan from PAK-HYM-1.0.

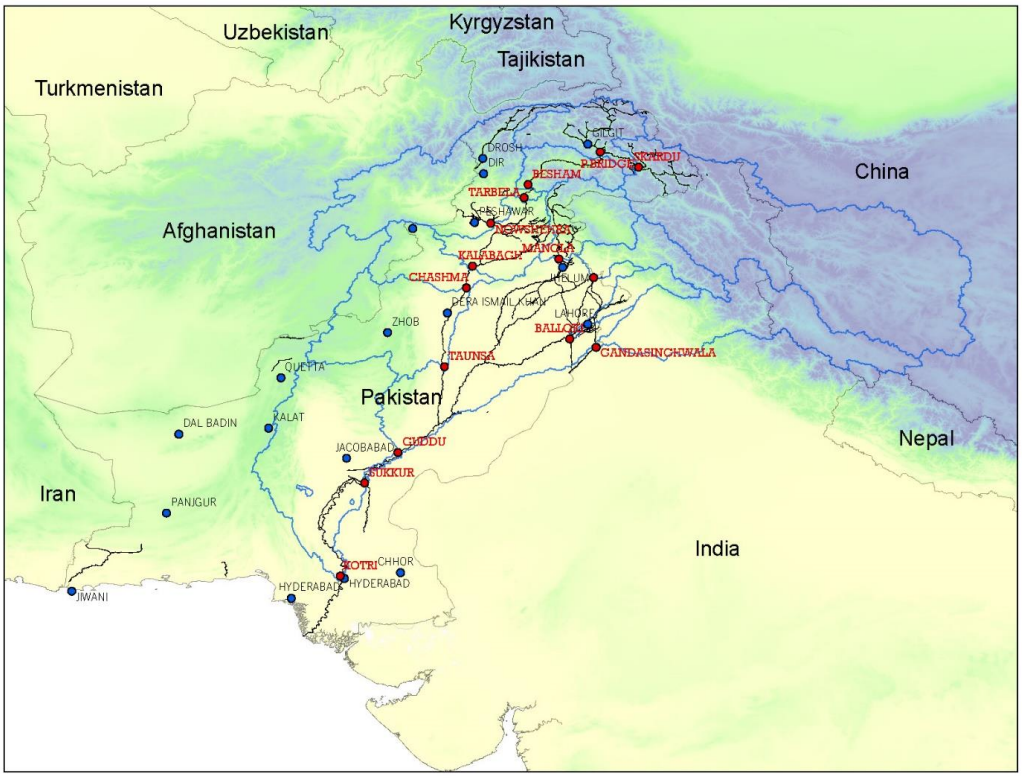


Figure C. 2. Map presenting Indus Basin (Blue boundaries) with some of the river inflows gauging locations (Red points) and long-term WMO affiliated weather stations (Blue points). Light color polygon indicates the political boundaries of Pakistan.

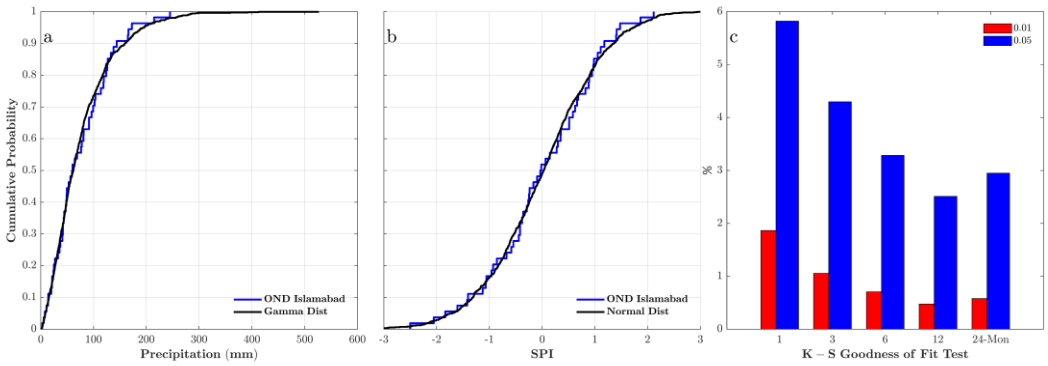


Figure C. 3. Example of equiprobability transformation from gamma distribution fit on precipitation to the standard normal distribution in SPI framework. Cumulative probability distribution of October-December (OND) Islamabad precipitation observations (blue line), fitted gamma distribution (black like) (a), similar to (a) with cumulative standard normal distribution (black line) in SPI framework (b), Rejection rates of the Kolmogorov–Smirnov goodness-of-fit test for gridded PAK-HYM-1.0 data for 1-month, 3-months, 6-months, 12-months and 24-months at significance levels of 0.01 (red bars) and 0.05 (blue bars).

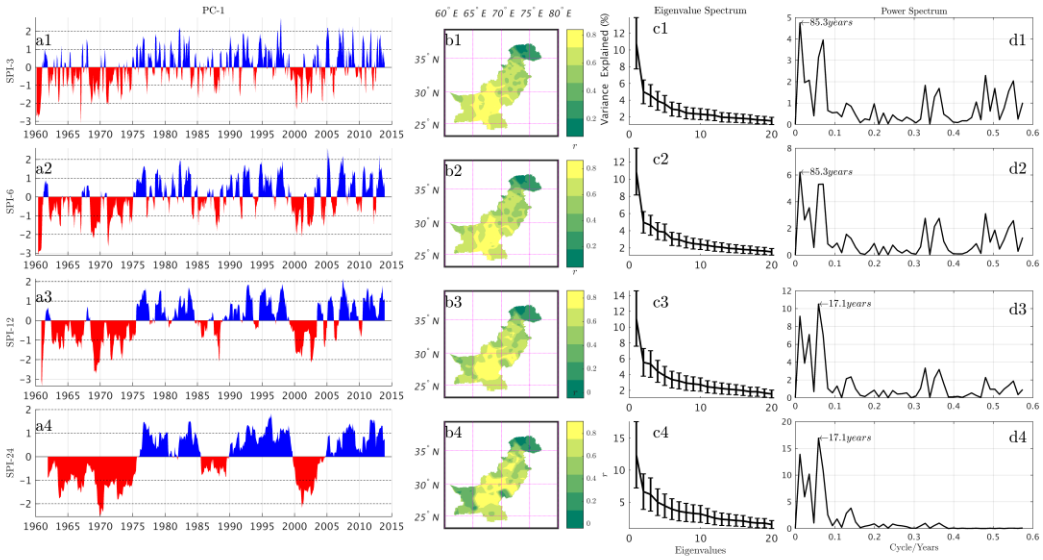


Figure C. 4. Analyses of Standardized Precipitation Index for Pakistan. PC-1 of the maximum covariance analysis for 3-months SPI is presented in (a1), where x-axis presents years and y-axis presents the magnitude of the standardized anomaly, blue color indicates positive anomalies and red color indicates negative anomalies, correlation map of PC-1 and 3-month SPI gridded time series, where color bar presents the strength of correlation (b1), Eigenvalue spectrum of covariance analysis of 3-month SPI with significance indicated by error bar (c1), periodogram of PC-1 of 3-month SPI covariance analysis indicating spectrum power density at y-axis and frequency at x-axis (d1). Similarly, 6-, 12- and 24- months analysis is presented in successive rows.

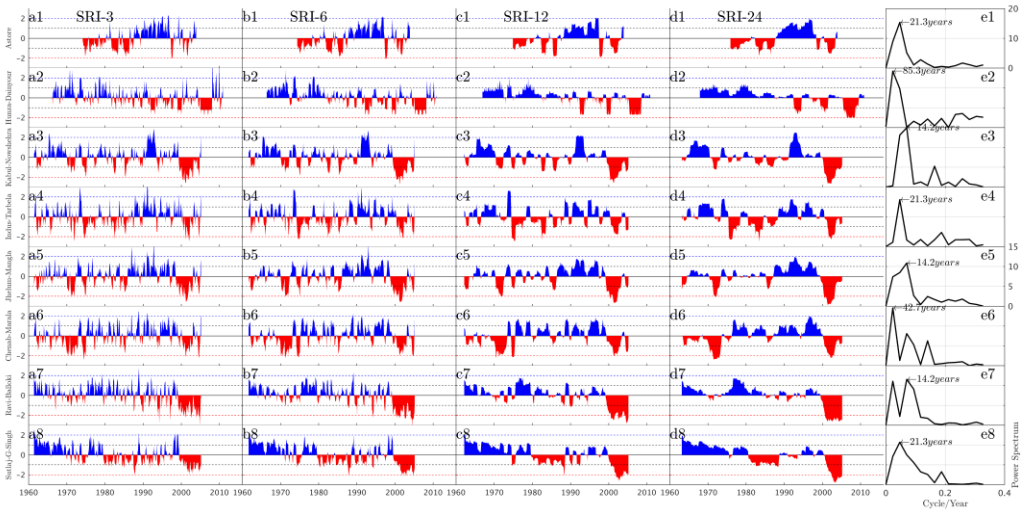


Figure C. 5. Standardized Runoff Index (SRI) for different rivers of Pakistan. 3-month SRI for River Astore (a1), 6-month SRI (a2), 12-month SRI (c1), 24-month SRI (d1) and spectral analysis indicating power density and frequency (e1). Same analyses for River Hunza at Dainyour (a2, b2, c2, d2, e2), for River Kabul at Nowshehra (a3, b3, c3, d3, e3), for River Indus at Tarbela (a4, b4, c4, d4, e4), for River Jhelum at Mangla (a5, b5, c5, d5, e5), for River Chenab at Marala (a6, b6, c6, d6, e6), for River Ravi at Balloki (a7, b7, c7, d7, e7), and for River Sutlaj at Qila Ganda Singhwala (a8, b8, c8, d8, e8). Panel e1, e2, e3 and e4 has similar y-axis presented in e1, whereas, e5, e6, e7 and e8 have similar y-axis presented in e5.

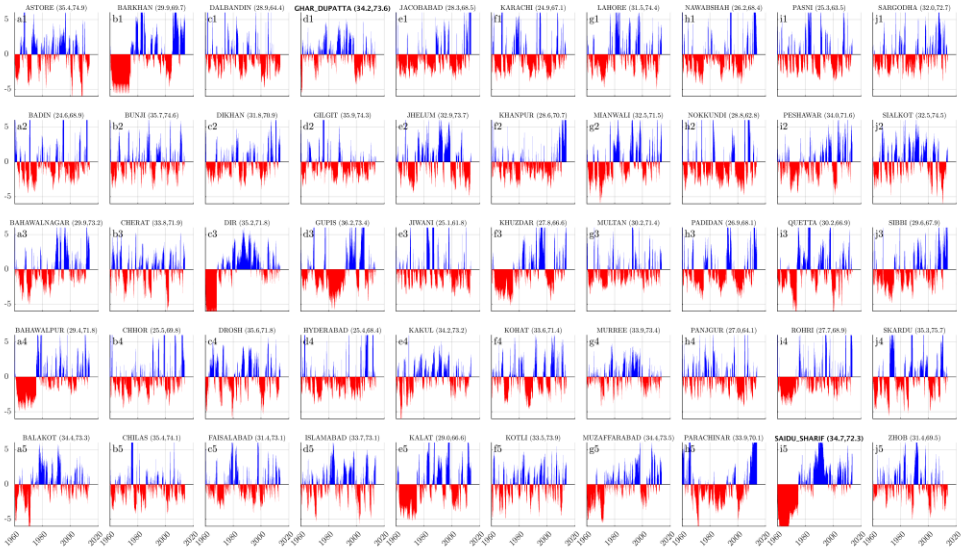


Figure C. 6. Palmer Drought Sensitivity Index (PDSI) for different location of Pakistan for the period of 1960-2013 is presented.

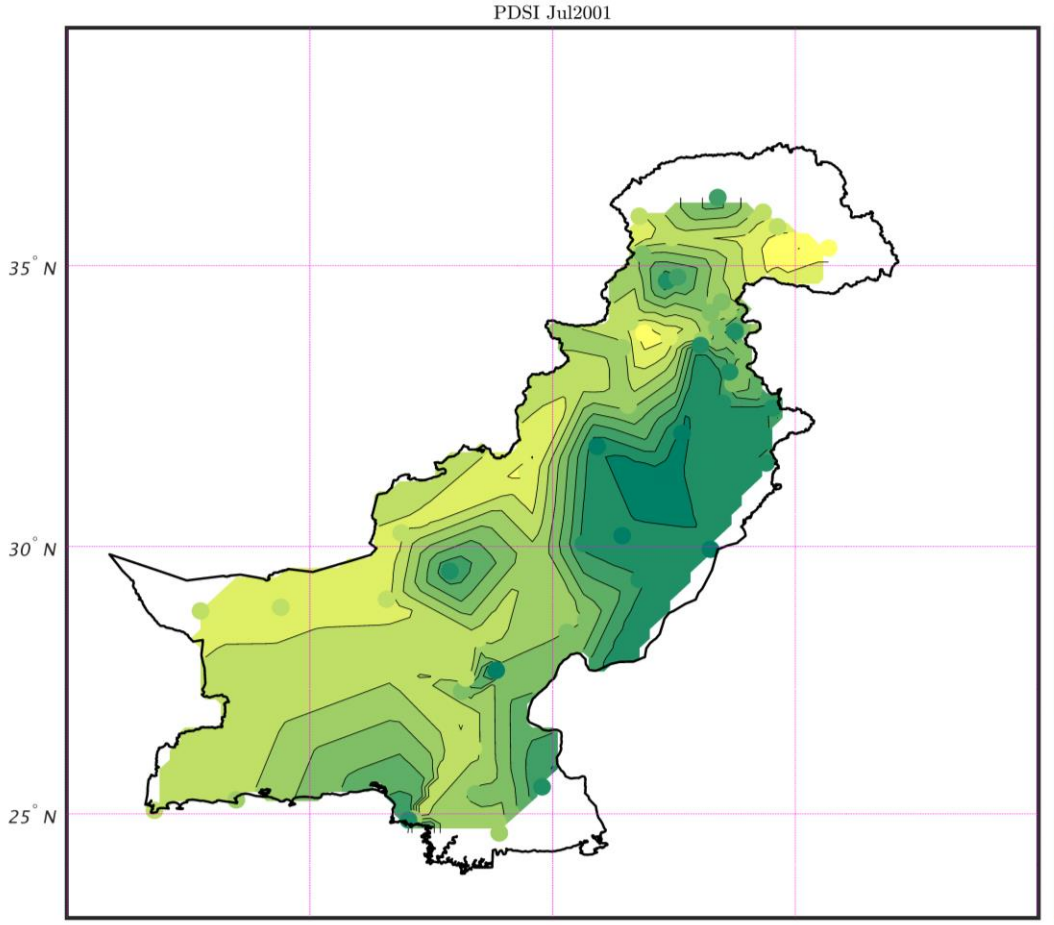


Figure C. 7. Palmer Drought Sensitivity Index map of Pakistan for July 2001.

SUPPORTING INFORMATION

AN INVESTIGATION AND ASSESSMENT OF METEOROLOGICAL AND HYDROLOGICAL DROUGHTS IN PAKISTAN USING HIGH RESOLUTION IN SITU GRIDDED DATA SET

SPI METHODOLOGY

Mckee *et al.* (1993) developed the Standardized Precipitation Index (SPI) for the purpose of defining and monitoring drought. It is one of the most popular drought monitoring index that is approved by almost all the National Meteorological and Hydrological Services across the world to characterize the meteorological drought. The SPI is a powerful, flexible index that is simple to calculate. In fact, precipitation is the only required input parameter. In addition, it allows an analyst to determine the rarity of a drought or an anomalously wet event at a particular time scale for any location in the world that has a precipitation record.

Thom (1966) found the utility of gamma distribution for the climatological precipitation time series. The gamma distribution is defined by its frequency or probability density function:

$$g(x) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1} e^{\frac{-x}{\beta}} \text{ for } x > 0 \quad \text{(S1)}$$

$$\alpha > 0$$
$$\alpha \text{ is a shape parameter} \quad \text{(S2)}$$

$$\beta > 0$$
$$\beta \text{ is a scale parameter} \quad \text{(S3)}$$

$$X > 0$$
$$X \text{ is the precipitation amount} \quad \text{(S4)}$$

$$\Gamma(\alpha) = \int_0^{\infty} y^{\alpha-1} e^{-y} dy$$
$$\Gamma(\alpha) \text{ is the gamma function} \quad \text{(S5)}$$

In general, the Gamma function is evaluated either numerically or using the values tabulated depending on the value taken by parameter α. In order to model the data observed with

a gamma distributed density function, it is necessary to estimate appropriately parameters α and β.

Different methods have been suggested in literature for the estimate of these parameters, for example, in Edwards and McKee (1997), the Thom (1958) approximation is used for the maximum probability

$$\hat{\alpha} = \frac{1}{4A}\left(1 + \sqrt{1 + \frac{4A}{3}}\right) \tag{S6}$$

$$\hat{\beta} = \frac{\bar{x}}{\hat{\alpha}} \tag{S7}$$

Where for n observations

$$\mathrm{A} = \ln(\bar{x}) - \frac{\sum \ln(x)}{n} \tag{S8}$$

After estimating coefficients α and β the density of probability function g(x) is integrated with respect to x and we obtain an expression for cumulative probability G(x) that a certain amount of rain has been observed for a given month and for a specific time scale.

$$G(x) = \int_0^x g(x)dx = \frac{1}{\hat{\beta}^{\hat{\alpha}}(\hat{\alpha})}\int_0^x x^{\hat{\alpha}-1} e^{\frac{-x}{\hat{\beta}}}\, dx \tag{S9}$$

Taking $t = \frac{x}{\hat{\beta}}$ this equation becomes the incomplete Gamma function

$$G(x) = \frac{1}{\Gamma(\hat{\alpha})}\int_0^x t^{\hat{\alpha}-1} e^{-t}\, dt \tag{S10}$$

The Gamma function is not defined by x = 0 and since there may be no precipitation the cumulative probability becomes:

$$H(x) = q + (1 - q)\, G(x) \tag{S11}$$

Where, q is the probability of no precipitation. The cumulative probability is then transformed into a normal standard distribution with null average and unit variance from which we obtain the SPI index.

We have used MATLAB routine to estimate the parameters of Gamma distribution and Gamma function over here.

Table S1 also contains the corresponding probabilities of occurrence for each level of severity arising naturally from the normal probability density function. Thus, at a given location for an individual month, moderate droughts (SPI≤−1) have a probability of occurrence 15.9%, whereas extreme droughts (SPI ≤−2) have a probability of 2.3%. The SPI will have extreme values, by definition, with the same frequency at all locations.

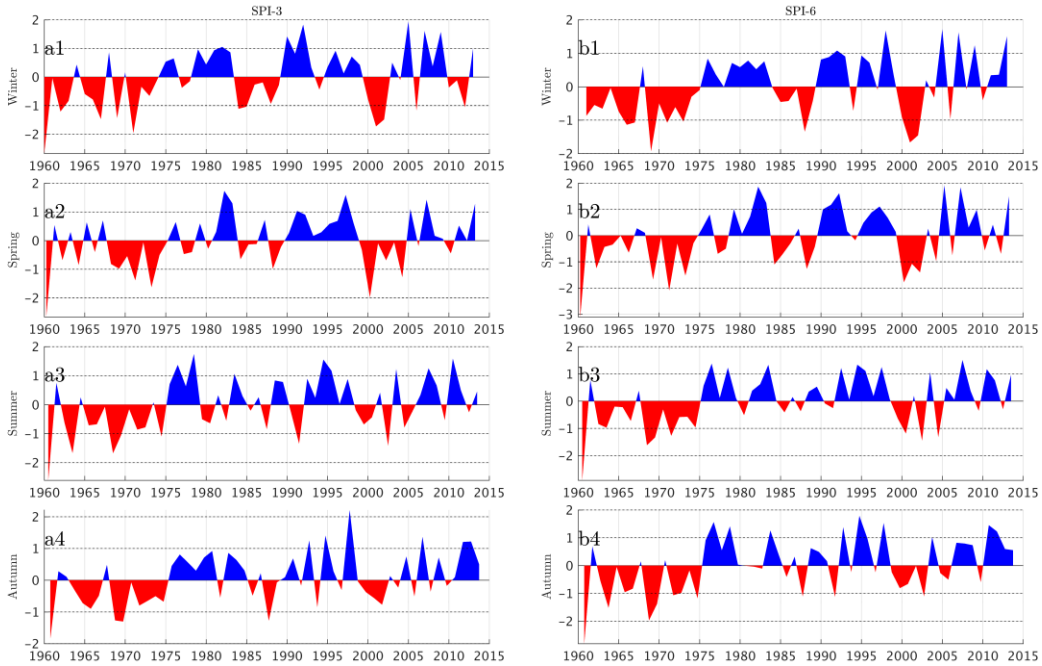


Figure S1. Seasonal decomposition of PC-1 of 3- and 6- month's gridded SPI. 3-month Winter SPI (a1), 3-month Spring SPI (a2), 3-month Summer SPI (a3), 3-month Autumn SPI (a4), 6-month Winter SPI (b1), 6-month Spring SPI (b2), 6-month Summer SPI (b3), 6-month Autumn SPI (b4).

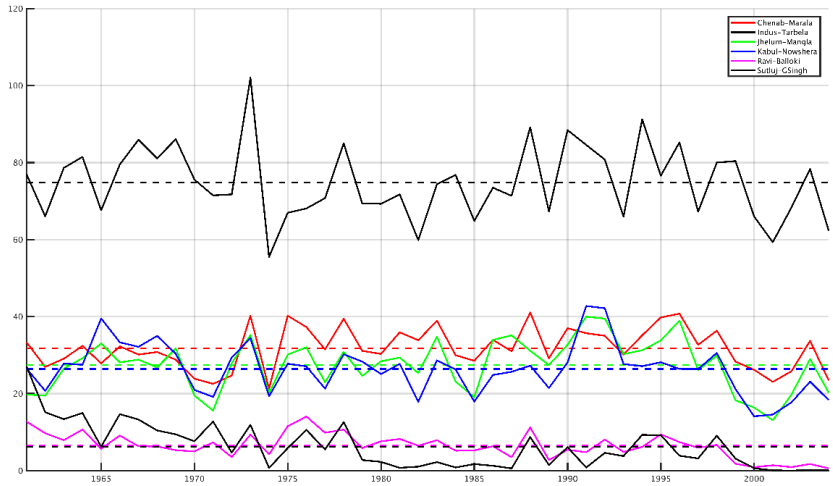


Figure S2. Annual cumulation of river inflows in different river of Pakistan from 1961-2005. Dash line indicates annual mean of accumulated water. X-axis presents years and y-axis presents volume of water in Km^3.

Table S1. Drought classification by SPI value and corresponding event probabilities

SPI value	**Category**	**Probability (%)**
2.00 or more	Extremely wet	2.3
1.50 to 1.99	Severely wet	4.4
1.00 to 1.49	Moderately wet	9.2
0 to 0.99	Mildly wet	34.1
0 to -0.99	Mild drought	34.1
-1.00 to -1.49	Moderate drought	9.2
-1.50 to -1.99	Severe drought	4.4
-2 or less	Extreme drought	2.3

Table S2. Drought Classification by PDSI values

PDSI Value	**Classification**
4.00 or more	Extremely wet
3.00 to 3.99	Very wet
2.00 to 2.99	Moderately wet
1.00 to 1.99	Slightly wet
0.50 to 0.99	Incipient wet spell
0.40 to -0.49	Near normal
-0.50 to -0.99	Incipient dry spell
-1.00 to -1.99	Mild drought
-2.00 to -2.99	Moderate drought
-3.00 to -3.99	Severe drought
-4 or less	Extreme drought