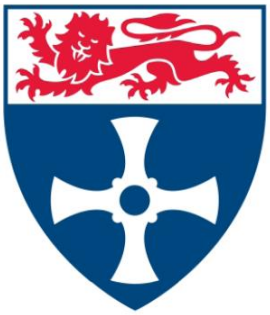
Evaluation of climate and hydrological models for impact projections in the Upper Indus Basin



David Pritchard

School of Engineering

Newcastle University

A thesis presented for the degree of Doctor of Philosophy

April 2019

Abstract

Water resources in the Indus basin are under acute and growing stress. How climate change will affect this situation in the coming decades depends substantially on responses in the data-sparse mountains of the upper basin. However, model projections of changes in the cryosphere-dominated hydrology here are highly uncertain. Integral to this uncertainty are challenges in: characterising near-surface climate fields needed for model input; selecting appropriate model structures to balance process fidelity with data availability; and understanding the wide spread in climate model projections used in impact assessments. As such, this thesis aims to identify pathways for refined hydrological projections in the upper Indus basin through in-depth evaluation of climate, cryospheric and hydrological models.

Firstly, using the High Asia Refined Analysis (HAR), the study assesses how relatively high resolution regional climate modelling can help describe spatiotemporal variability in near-surface climate. The HAR exhibits substantial skill in many respects, but particularly in capturing the complex patterns of precipitation in the basin. Some seasonally varying biases in temperature and incoming radiation suggest deficiencies in snow and cloud representations that are likely resolvable. Secondly, the Factorial Snowpack Model (FSM) is driven with the HAR to examine the feasibility and required structure of process-based snowpack modelling. Model correspondence with local observations and remote sensing is good for a subset of FSM configurations using a prognostic albedo parameterisation, as well as a representation of liquid water retention, drainage and melt/refreezing cycles in the snowpack. The multi-physics approach additionally highlights the inputs and processes needing further investigation, which include the atmospheric stability adjustment. Thirdly, using an adapted FSM program and TOPKAPI-ETH, simplified representations of cryospheric processes are compared with more process-based approaches. This helps to identify where systematic differences in hydrological response occur and their connection with spatial and temporal scales. It is found that an enhanced temperature index (ETI) model exhibits behaviour and climate sensitivity more akin to energy balance formulations than a classical temperature index model. However, there may be structural limits to the fidelity of the ETI formulation under cloudy conditions, while further attention is needed on the translation of surface melt to runoff, especially at high elevations.

The study then moves to examine controls on regional trends and variability simulated by climate models, focusing on temperature in CMIP5 GCMs. While the models partly reproduce key regional atmospheric circulation influences, variation in summer temperature responses depends on differing snow and albedo representations. Ultimately this may offer some potential to constrain temperature projections. Finally, using CMIP5 and HAPPI GCM outputs, the study explores climate and hydrological projections under selected global warming stabilisation scenarios. This shows that shifts in the timing of runoff are discernible even for low warming targets. Overall water availability may depend particularly on natural variability in precipitation, but in dry years the pressures on water resources in the basin could worsen in future. Further efforts to constrain the range of projections using observations and process-based reasoning are required, but effective water resources management in the basin is likely to depend on increasing resilience to a wide range of climatic and hydrological variability.

Acknowledgements

I would like to thank my supervisors Hayley Fowler, Nathan Forsythe and Greg O'Donnell for all of their guidance, help and support during this project. Your ideas, insights and patience have all been very much appreciated. Thanks are also due to Nick Rutter for sharing his expertise on snow processes and modelling, as well as András Bárdossy for his ideas during the first part of the project. I am also most grateful to the many organisations and individuals whose data collection, data product development and model code have underpinned my research. In this respect, particular thanks are due to Fabien Maussion and Richard Essery for putting their very valuable work into the public domain, as well as Francesca Pellicciotti for sharing the TOPKAPI-ETH model. I also gratefully acknowledge the efforts made by the many people involved in collecting local observations in the Indus basin, creating global remote sensing and reanalysis products, and producing climate model simulations. Finally I would like to thank my family for always being there and Liz for being so patient, encouraging and supportive throughout this process.

Contents

List of Publications

A version of Chapter 3 has been published in the Journal of Hydrometeorology. I conducted all of the data analysis, figure development and writing, with the co-authors contributing suggestions and some manuscript editing.

Pritchard, D.M.W., N. Forsythe, H. J. Fowler, G. O'Donnell and X.-F. Li, 2019: Evaluation of Upper Indus near-surface climate representation by WRF in the High Asia Refined Analysis. *J. Hydrometeorol.*, **20**, 467–487, doi:10.1175/JHM-D-18-0030.1.

I also co-authored two publications while undertaking this project. These publications are not part of the thesis, but the work presented in Chapter 6 follows on from both manuscripts.

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**, 664–670, doi:10.1038/NCLIMATE3361.

Li, X.-F., H. Fowler, N. Forsythe, S. Blenkinsop, and D. Pritchard, 2018: The Karakoram/Western Tibetan vortex: seasonal and year-to-year variability. *Clim. Dyn.*, 1–24, doi:10.1007/s00382-018-4118-2.

List of Figures

List of Tables

List of Abbreviations

AAR	Accumulation Area Ratio
AMIP	Atmospheric Model Intercomparison Project
APHRODITE	Asian Precipitation Highly Resolved Observational Data Integration Towards Water Resources
AR5	Fifth Assessment Report of the Intergovernmental Panel on Climate Change
AVHRR	Advanced Very High Resolution Radiometer
AWS	Automatic Weather Station
CCF	Cloud Cover Fraction
CFSR	Climate Forecast System Reanalysis
CGT	Circumglobal Teleconnection
CMA	Continuous Melt Area
CMIP5	Coupled Model Intercomparison Project Phase 5
DDF	Degree Day Factor
DEB	Debris Energy Balance (model)
DEM	Digital Elevation Model
DETI	Debris Enhanced Temperature Index (model)
DTR	Diurnal Temperature Range
EB	Energy Balance
ELA	Equilibrium Line Altitude
ENSO	El Niño Southern Oscillation
ERAI	ERA-Interim
ESA	European Space Agency
ET	Evapotranspiration
ET0	Reference Evapotranspiration
ETI	Enhanced Temperature Index (model)
EvK2CNR	Everest-K2-Consiglio Nazionale delle Ricerche (research group)
FAO	Food and Agriculture Organisation of the United Nations
FNL	Final Operational Global Analysis (dataset)
FSM	Factorial Snowpack Model
FSM+	Factorial Snowpack Model (with modifications introduced in this work)
GCM	General Circulation Model
GFS	Global Forecast System

GOP	Government of Pakistan
HAPPI	Half a degree Additional warming, Prognosis and Projected Impacts
HAR	High Asia Refined Analysis
HAR10	High Asia Refined Analysis (10km resolution domain)
HAR30	High Asia Refined Analysis (30km resolution domain)
HKKH	Hindu Kush-Karakoram-Himalaya
ICESat	Ice, Cloud and land Elevation Satellite
IPCC	Intergovernmental Panel on Climate Change
JRA55	Japanese 55-year Reanalysis
KV	Karakoram Vortex
KZI	Karakoram Zonal Index
KZS	Karakoram Zonal Shear
LSM	Land Surface Model
LST	Land Surface Temperature
mASL	Metres Above Sea Level
MERRA	Modern-Era Retrospective analysis for Research and Applications
MERRA2	Modern-Era Retrospective analysis for Research and Applications, Version 2
MJO	Madden-Julian Oscillation
MMM	Multi-Model Mean
MODIS	Moderate Resolution Imaging Spectroradiometer
MYJ	Mellor-Yamada-Janjić (planetary boundary layer scheme)
NAO	North Atlantic Oscillation
NASA	National Aeronautics and Space Administration
NCEP	National Centers for Environmental Prediction
NDSI	Normalised Difference Snow Index
NWP	Numerical Weather Prediction
NWUIB	North-West Upper Indus Basin
PBL	Planetary Boundary Layer
PMD	Pakistan Meteorological Department
RCM	Regional Climate Model
RCP	Representative Concentration Pathway
RGI	Randolph Glacier Inventory
RMSE	Root Mean Square Error
SASM	South Asian Summer Monsoon

SCA	Snow-Covered Area
SEB	Simplified Energy Balance
SMAP	Soil Moisture Active Passive (NASA project)
SnowMIP	Snow Model Intercomparison Project
SnowMIP2	Phase 2 of the Snow Model Intercomparison Project
SRF	Shortwave Radiation Factor
SRTM	Shuttle Radar Topography Mission
SST	Sea Surface Temperature
SWE	Snow Water Equivalent
SWJ	Subtropical Westerly Jet
T2	Air Temperature at 2m
T500	Air Temperature at 500 hPa
TF	Temperature Factor
TI	Temperature Index (model)
TOPKAPI	TOPographic Kinematic Approximation and Integration (model)
TOPKAPI-ETH	TOPographic Kinematic Approximation and Integration (model, Eidgenössische Technische Hochschule Zürich variant)
TOPMODEL	TOPography-based hydrological MODEL
TRMM	Tropical Rainfall Measuring Mission
Tt	Temperature Threshold
UIB	Upper Indus Basin
WAPDA	Water and Power Development Authority
WRF	Weather Research and Forecasting (model)
WYMI	Webster Yang Monsoon Index

Chapter 1

Introduction

1.1 Background

Water resources in the Indus basin are under acute and growing stress. In Pakistan, whose borders contain 63% of the basin, annual water availability per capita declined from approximately 5000 m^3 in 1950 to 1000 m^3 by 2010 (Briscoe 2010; Karki et al. 2011; UNDP 2016). This stark reduction means the country is now hovering around the threshold used to define water scarcity (UNDP 2006). Such scarcity at the national scale cannot be ameliorated simply with supply-side measures. The available sources of water are in large part fully allocated for human uses, with the vast majority reserved for the irrigated agriculture so vital to livelihoods and the economy (Briscoe and Qamar 2007; Archer et al. 2010). This means that providing sufficient water to users in drought years is a particular problem. Only around 30 days' worth of the annual flow can be stored in the basin's ageing infrastructure to alleviate dry conditions, compared with 900 days in the Colorado River basin, for example (Condon et al. 2014). The situation worsens when we consider that the substantial rates of groundwater abstraction in the low-lying plains are unsustainable (Sattar et al. 2017).

It is therefore a major issue that the demand for water continues to accelerate. One driver is the nation's immense population growth (Archer et al. 2010). Following a fivefold increase since 1950, the population is projected to rise from around 197 million in 2017 to 307 million by 2050 (UN 2017). Rapid urbanisation and industrialisation compound the implications of population growth alone, such that shortfalls in water supply are projected for the coming decades (Zawahri and Michel 2018). This is before even considering provisions for the environmental flows needed to support the Indus delta and other sensitive ecosystems (Inam et al. 2007). As such, water security seems liable to deteriorate further for the country as a whole. This burden will continue to be unequally distributed amongst the population, owing to the social, political and historical processes that produce uneven structures and realisations of scarcity (Akhter 2017).

In conjunction with other issues, including deteriorating water and soil quality in vast agricultural areas, as well as acute flooding events, these challenges are termed an "emerging

water crisis" in the National Water Policy produced by the Government of Pakistan in April 2018 (GOP 2018, p. 1). Within a framework of integrated catchment management, the Policy identifies a number of objectives to resolve these issues. Strategic priorities include water conservation and efficiency measures at multiple scales and across sectors. There is also an intention to develop storage options to buffer hydroclimatic variability and better harness the basin's considerable renewable energy potential. However, the challenges to the Policy's implementation are formidable. Relationships between the complex mosaic of institutions with water management responsibilities are often complicated and subject to bureaucratic inertia (Wescoat et al. 1991; Sattar et al. 2017). There are also substantial tensions between the nation's provinces, as exemplified by the history of disagreements over the 1991 Water Apportionment Accord (Condon et al. 2014; Anwar and Bhatti 2018). As a transnational basin, successful management is additionally in thrall to the volatile relationships between nations sharing the resources (Briscoe 2010; Zawahri and Michel 2018).

Compounding these multiple and interacting challenges is climate change. Efforts to unravel the potential impacts on the basin's hydrology have intensified in recent years, but uncertainties in how river flow magnitude, timing, variability and extremes might change in the coming years are still very large (Briscoe and Qamar 2007; Archer et al. 2010; Condon et al. 2014; Sattar et al. 2017; Bolch et al. 2012; Lutz et al. 2016a). Of particular concern are climate change impacts on runoff from the high mountains of the upper basin (hereafter Upper Indus Basin, UIB – Figure 1.1). As much of the Indus basin experiences low precipitation, the high-yielding catchments of the Hindu Kush, Karakoram and western Himalayan mountain ranges are vital for the water supplies, irrigated agriculture and hydropower on which downstream populations depend (Archer et al. 2010). Much of this runoff is derived from snow and glacier melt, supported in part by the largest perennial ice mass outside the poles. As the effects of climate change may be particularly pronounced at high elevations and in cryosphere-dominated runoff regimes (Pepin et al. 2015; Barnett et al. 2005), understanding the functioning and fate of the coupled climatic, cryospheric and hydrological system in the mountains is therefore paramount. A full description of the UIB and its hydroclimatology is given in the next chapter (Section 2.2).

In line with Wilby and Dessai's (2010) conceptualisation, uncertainty in UIB hydrological projections cascades and grows with each successive stage of the climate change impact assessment process. This process generally starts by defining emissions and mitigation scenarios, which is followed by climate modelling, regional downscaling and hydrological

impact modelling. Each component of the chain is beset by substantial complexity, but we contend that there are three challenges pertaining to climate and hydrological modelling that require particularly urgent attention in the data-sparse UIB. Firstly, refined near-surface climate characterisation is needed to improve hydrological model inputs. Secondly, comparison of alternative hydrological modelling approaches is required to identify those model structures balancing process fidelity and data availability. Thirdly, further evaluation of climate models is needed to understand the drivers of the wide spread in hydrological projections. Expanded in the next section, these challenges guide the design of this research.

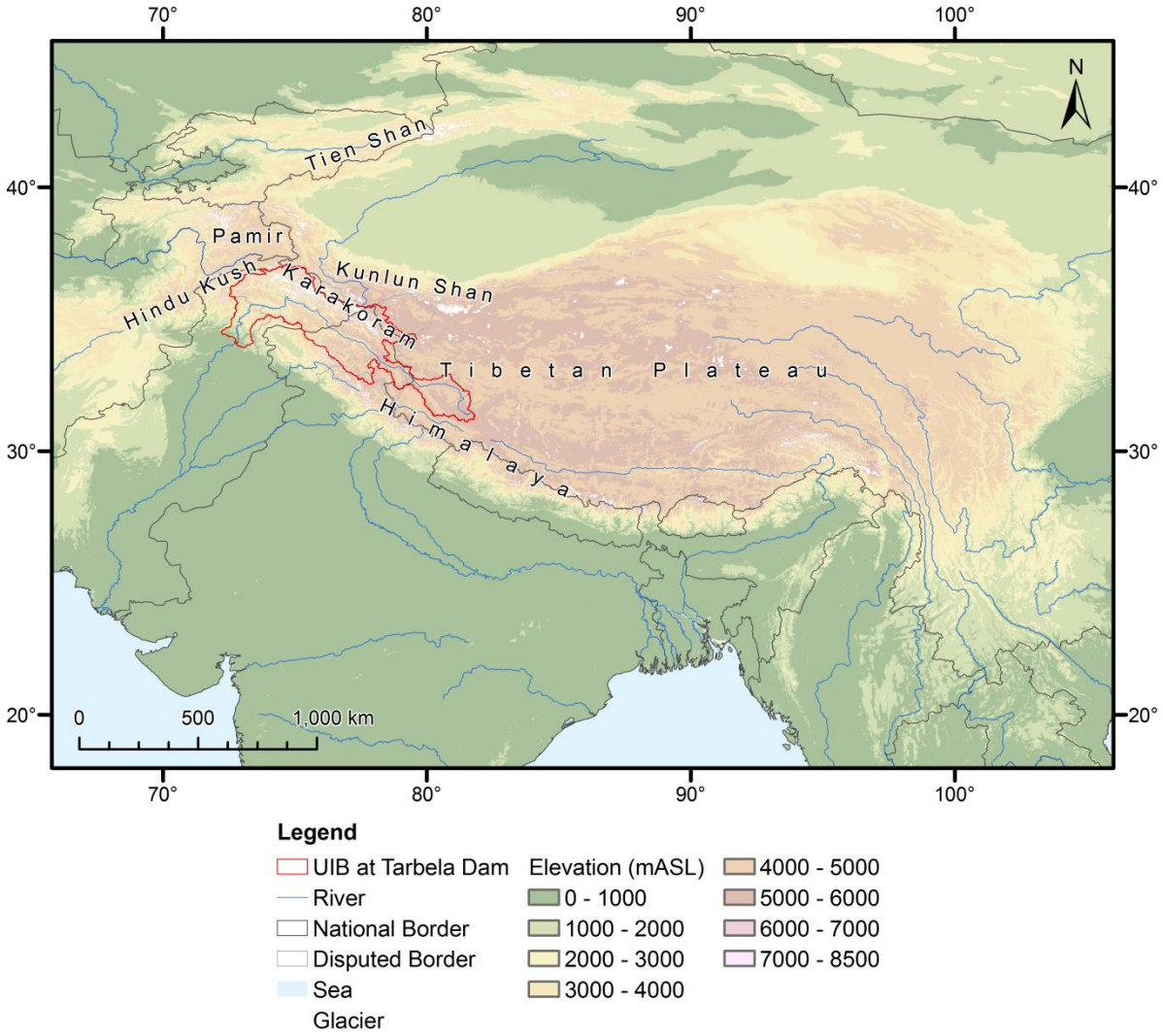


Figure 1.1 – Map of the Upper Indus basin and surrounding region.

1.2 Research Needs

1.2.1 Near-Surface Climate Characterisation

Nearly 30 years ago, Klemeš (1990) described how the intersection of sparse observation networks with pronounced climate variability is at the core of problems in modelling mountain hydrology. While many advances have been made in the intervening period, this data paucity continues to present obstacles in UIB modelling. This is exemplified by issues in attaining credible water balance closure (see e.g. Immerzeel et al. 2009, 2010; Reggiani and Rientjes 2015). Some recent studies have partly addressed this by calibrating precipitation inputs to obtain agreement with observed river flows and/or glacial mass balance (e.g. Immerzeel et al. 2012b, 2015; Ragettli et al. 2013). The approach certainly represents an advance over previous work, but it is subject to uncertainties in calibration data and model configurations, while typically simplifying spatial and temporal variability. More generally, it is inherently difficult to develop physically consistent climate fields for multiple variables at high spatial and temporal resolution from sparse observations alone. These fields are particularly required to test the feasibility of process-based hydrological modelling strategies, as discussed below.

One possible response to this problem harnesses modern computing facilities to simulate near-surface climate using high resolution numerical weather models (NWPs) or regional climate models. For example, it has been shown that the Weather Research and Forecasting (WRF) model (Skamarock et al. 2008) can perform well for hydrological applications in mountain regions, despite the complexity of terrain and the variability of climate processes (Ikeda et al. 2010; Rasmussen et al. 2011; Prein et al. 2012; Silverman et al. 2013; Duethmann et al. 2013). The small but growing number of applications in the Hindu Kush, Karakoram and Himalayan ranges show similar potential (e.g. Collier et al. 2013, 2015; Norris et al. 2016). However, application and evaluation of this approach for climate characterisation in the UIB has been very limited. This is particularly so at the basin and sub-basin scales of hydrological models, as well as over multi-year periods. As such, there is a need to assess how well such relatively high resolution regional climate modelling can augment other datasets and provide improved spatiotemporal climate fields for the basin. To explore this, we focus on the publicly available WRF simulations comprising the High Asia Refined Analysis (HAR) (Maussion et al. 2014), which has not been evaluated or applied in the UIB.

1.2.2 Hydrological Model Structures

Klemeš (1990) also discussed the difficulties in identifying appropriate model structures and parameterisations for the array of hydrological processes operating in mountain environments. In the UIB this issue is bound up with the particularly substantial limitations afflicting input and evaluation data noted previously. Quite understandably, it has led to the proliferation of relatively simple hydrological modelling approaches in the basin, as well as across the Himalayan arc more generally (e.g. Bocchiola et al. 2011; Rees and Collins 2006; Immerzeel et al. 2009, 2012a, 2013; Soncini et al. 2014; Ragettli et al. 2013, 2015; Immerzeel et al. 2014; Minora et al. 2015). As such, there have been few applications of process-based, energy balance models for hydrological applications here at all (but see e.g. Shrestha et al. 2015; Brown et al. 2014; Prasch et al. 2013). Yet these models may confer advantages for testing hypotheses about processes, completing mass and energy balance frameworks in coupled land-atmosphere models, and making projections in non-stationary climates. The question is thus whether application of such models is now feasible and useful in UIB hydrological studies. This depends in large part on the success of deriving climate input fields from regional climate models, as outlined above.

With this in mind, recent developments in systematic modelling frameworks present an interesting way to help assess the skill of such process-based modelling in the UIB. The framework approach usefully removes the confounding influences of variations in implementations between programs (Clark et al. 2015a,b). As such, it becomes possible to neatly synthesise the array of models in existence, while also more precisely delineating variations in model response that arise from alternative process representations and modelling decisions. Such frameworks have not been deployed so far in the Himalayan region, largely due to the difficulty in applying typical methods of performance evaluation with limited data. Yet this alone does not negate the need to better understand the implications of model configuration choices for practical applications. It also does not preclude the possibility of modelling frameworks being used for improving conceptual understanding of relevant processes, identifying critical areas for further investigation or taking an ensemble-based approach to explore uncertainties.

One framework of particular interest is the Factorial Snowpack Model (FSM) (Essery 2015). As an intermediate complexity multi-physics ensemble, FSM provides a means of testing alternative representations of key snowpack processes and how they interact with each other

within a coupled mass and energy balance scheme. Not only does this form a systematic framework of the kind discussed above, it also usefully allows us to place an emphasis on snow. This is critical, as physical snow processes have been little studied in the Himalayan region, despite their vast importance. For example, snow melt is the single largest component of river flow in the UIB (Mukhopadhyay and Khan 2015). Snow processes also strongly influence regional climate, through land-atmosphere interactions, and glacial mass balance. The patterns and processes underlying variability and changes in seasonal snow cover are therefore critical for both climate and hydrological projections. The questions that follow from this pertain to the combined skill of new input datasets and FSM, as well as the inferences that can be made about model configuration choices. Addressing these points necessitates enhanced use of remote sensing for evaluation at the level of individual processes, as far as possible, in order to augment the constraint on mass balance provided by observed river flows.

1.2.3 Climate Models and Projections

Previous studies have evaluated a number of pertinent aspects of climate model behaviour over the Himalaya and Tibetan Plateau. This includes the climatology of key variables (e.g. Su et al. 2013; Palazzi et al. 2015), near-surface processes (e.g. Rangwala et al. 2016; Palazzi et al. 2017) and aspects of regional circulation (e.g. Sperber et al. 2013; Levine et al. 2013). However, uncertainty in regional climate and hydrological projections remains large, due to substantial inter-model variation (e.g. Ragettli et al. 2013; Lutz et al. 2014, 2016a,b; Palazzi et al. 2015). It is therefore timely to consider how climate models simulate some of the processes underlying patterns of variability and change in critical near-surface variables.

In this respect, one priority is to assess how well climate models simulate the Karakoram Vortex (KV) regional atmospheric circulation patterns (Forsythe et al. 2017; Li et al. 2018). In brief, the KV describes anomalous circulation patterns centred on or near the Karakoram throughout the year. Anticyclonic and cyclonic KV states are closely linked to anomalies in the position and intensity of the subtropical westerly jet (SWJ) in all seasons, as well as the South Asian monsoon (SASM) in summer. The strong, spatially extensive and homogenous temperature response to KV circulation anomalies in winter contracts substantially in summer, becoming focused on a more restricted zone around the Karakoram. In conjunction with trends in the SWJ, SASM and KV states, Forsythe et al. (2017) argued that this contraction during the peak ablation season provides a significant contribution to the unique

summer cooling (Fowler and Archer 2006; Forsythe et al. 2012b), neutral glacier mass balance (e.g. Hewitt 2005; Zhou et al. 2017; Bolch et al. 2017; Brun et al. 2017), and stable or declining river flow trends (Sharif et al. 2013) observed in the heavily glaciated catchments of the Karakoram and UIB. Although this is an emerging area of research, similar findings were recently reached by Mölg et al. (2017) and Norris et al. (2018). As such, understanding how well the KV system is represented by climate models may be critical for gauging the plausibility of the climate trajectories they simulate.

With respect to these trajectories, a notable gap exists in that hydrological projections for the UIB have not yet explicitly considered possible changes in river flow regimes under the climate change targets set out in the Paris Agreement of 2015. These targets aim to limit global temperature change to 2°C above pre-industrial levels, with endeavours to restrict the change to 1.5°C. More in line with current trajectories, we could also ask what the implications of 3 or 4°C warmer worlds may be. Rather than characterising how such stabilisation scenarios may play out in the UIB, most previous work in the basin has focused on fixed time slices in the future (e.g. Immerzeel et al. 2013; Lutz et al. 2014; Soncini et al. 2014; Ali et al. 2015). As such, framing the problem in terms of policy-relevant stabilisation scenarios raises some interesting questions about how much projections might vary for the same warming scenario in models of differing climate sensitivity, as well as appropriate methodology to identify any forced response in a context of substantial internal variability. There is thus a need to explore hydrological projections for these scenarios, in order to see if any clear signals are likely to emerge that could ultimately guide adaptation and support the broader evidence base motivating mitigation efforts.

1.3 Aim and Objectives

In light of the research needs outlined above, the aim of this thesis is to identify pathways for refined hydrological projections in the upper Indus basin. The thesis is structured around the following objectives:

1. To evaluate how well the HAR represents near-surface climate in the UIB, with reference to its potential for improving climate inputs for hydrological modelling

2. To assess the feasibility of process-based snow modelling in the UIB, as well as the implications of alternative process representations

3. To compare process-based models of snow and glacier ablation with simpler approaches for simulating UIB hydrology

4. To analyse climate model skill in representing key controls on temperature variability and projections, focusing on the Karakoram Vortex

5. To explore the implications of selected global warming targets for the climate and hydrology of the UIB.

1.4 Thesis Structure

The remainder of the thesis begins with an introduction to the UIB and a review of research on its climate, cryosphere and hydrology in Chapter 2. This helps to introduce some key background information on the UIB required to understand the subsequent chapters. Chapter 2 also contains some further discussion of the research needs outlined above, although each of the subsequent chapters has a fairly detailed introduction taking the relevant literature into account.

Each of the subsequent chapters focuses on one of the research objectives. The relationships between the chapters are summarised schematically in Figure 1.2. Chapter 3 evaluates how well the HAR simulates near-surface climate in the UIB, while Chapter 4 deploys the HAR in process-based snow modelling using FSM. The emphasis here is on establishing the feasibility of the input/model combination, as well as delineating the effects of different parameterisation options within a multi-physics ensemble. Chapter 5 introduces the extensions to this ensemble developed as part of this thesis for simulating key snow and glacier processes in the UIB. The extended version of FSM (referred to here as FSM+) is used alongside TOPKAPI-ETH in Chapter 5 to compare and evaluate cryospheric and hydrological modelling approaches of varying complexities. Climate model performance and behaviour are then evaluated in Chapter 6, focusing on critical controls on simulated temperature variability and trends. Chapter 7 then explores hydrological projections under selected climate stabilisation scenarios. Conclusions and further discussion are provided in Chapter 8.

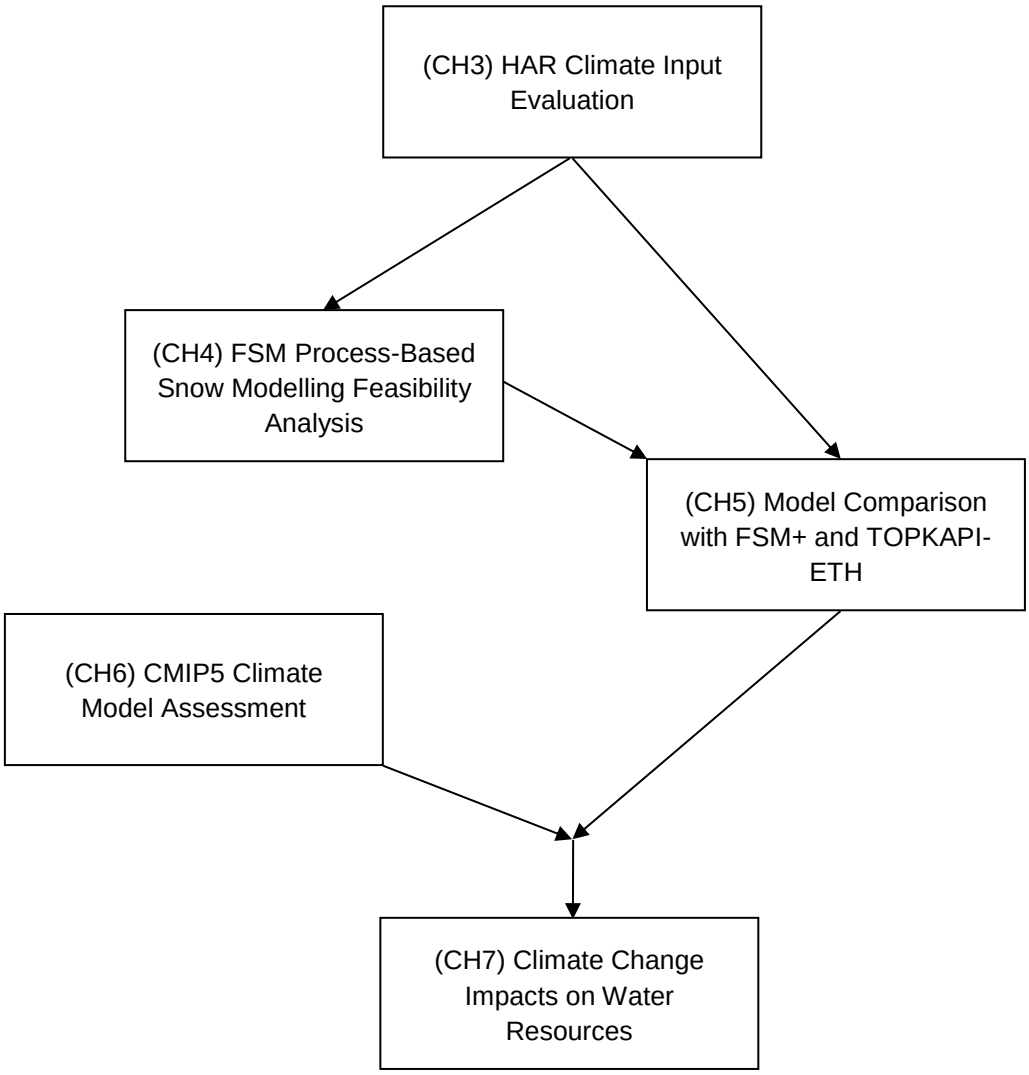


Figure 1.2 – Schematic representation of thesis structure and relationships between chapters 3 to 7.

Chapter 2

Review of Climate and Hydrology Research in the Upper Indus Basin

2.1 Introduction

This chapter reviews previous research on the climate and hydrology of the Upper Indus Basin (UIB). Substantial emphasis is placed on reviewing the literature relevant to our understanding of the behaviour of the coupled climatic, cryospheric and hydrological systems in the basin. This provides the context required to understand the results presented in subsequent chapters. An overview of the literature supporting the research gaps identified in Chapter 1 is also provided. However, more detailed discussion of the relevant studies is provided at the beginning of Chapters 3 to 7.

2.2 Background on Upper Indus Basin

Before reviewing our current understanding of the climate, cryosphere and hydrology of the UIB, this section provides some background information on the basin. Figure 2.1 shows a map of the UIB and its river network, gauged sub-basins, glaciated areas and climate station locations. While around 11% of the UIB area is glaciated according to the Randolph Glacier Inventory (RGI) 5.0 (Arendt et al. 2015), Figure 2.1 demonstrates that the majority of the UIB's glaciers are situated in the Hunza, Shigar and Shyok sub-basins, which drain the Karakoram range (Figure 1.1). The very long tongues of some Karakoram glaciers are visible, with several glaciers running to tens of kilometres in length. The larger glaciated fractions of the three sub-basins draining the Karakoram range are accompanied by differences in hypsometry compared with the overall UIB, as reflected in Figure 2.2. This shows a shift upwards to higher elevations and a sharpening in the hypsometric curve for the Karakoram sub-basins relative to the whole UIB. Peaking at a higher elevation again, the larger glacial fraction in the Karakoram sub-basins is also visible on Figure 2.2.

Based on the ESA GlobCover 2009 land cover product (Arino et al. 2012), 73% of the non-glaciated area of the basin is comprised of a mixture of bare ground and herbaceous plants, such as grasses and low shrubs. Much of the remaining area is classified as mixed vegetation and cropland, which is found predominantly in the lower reaches of the UIB, around and upstream of Tarbela Dam. While 11% of the UIB is classified as perennial cryosphere through

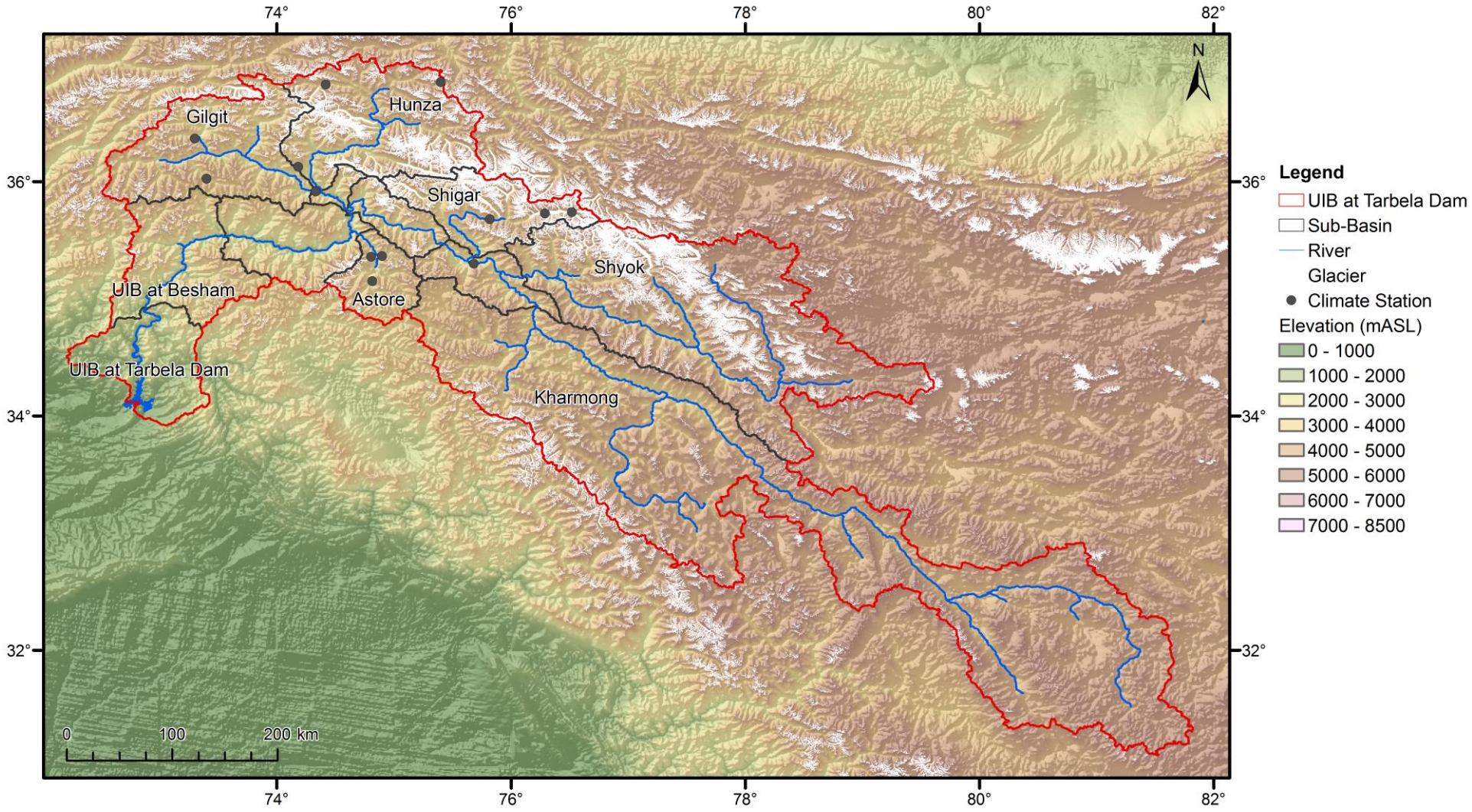


Figure 2.1 – Map of the Upper Indus Basin showing the river network, gauged sub-basins (key ones labelled), glaciated areas and climate station locations.

glacier cover, the majority of the basin is subject to seasonal snow cover. Using MODIS remote sensing, Forsythe et al. (2012b) demonstrated that around 70% of the UIB is snow-covered during the winter peak of the annual snow cover cycle. In some of the sub-basins, including Astore and Gilgit, the maximum annual snow cover routinely exceeds 90%. Seasonal snow cover thus has large implications for the climate and hydrology of the UIB.

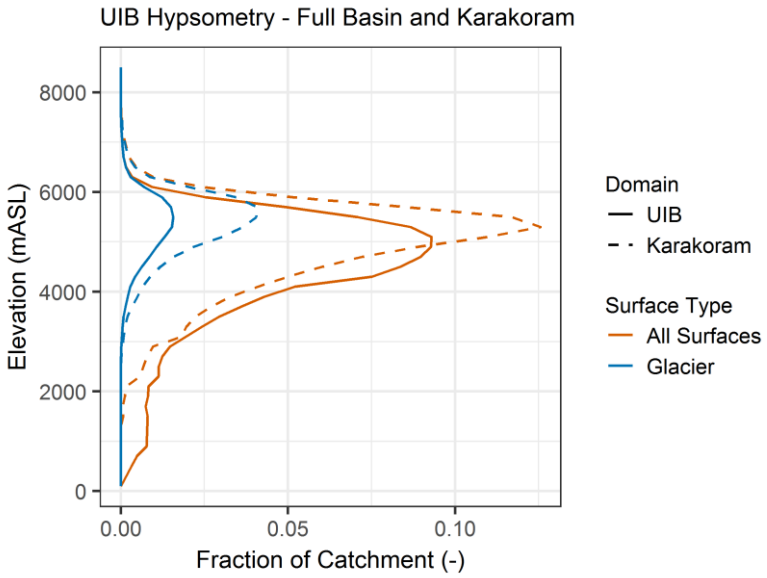


Figure 2.2 – UIB hypsometry showing the full basin (solid lines) and Karakoram (dashed lines), with the latter based on the Hunza, Shigar and Shyok sub-basins. Curves for all surface types (orange) and glacier surfaces only (blue) are plotted.

Figure 2.1 also reflects the sparsity of the climate observation network, along with the bias of measurements towards valley locations in the north-west part of the basin. Only three of the stations have long records extending back more than 30 years (Section 3.3.2). In contrast, many of the sub-basins shown in Figure 2.1 have fairly long river flow records beginning in the 1960s or 1970s (Section 3.3.2). The smallest gauged sub-basin is the Astore, which covers an area of 3988 km^2, while the UIB at Tarbela Dam is 172173 km^2. Human influences and alterations in the basin are fairly minor, although the settlements that are present in the UIB have to contend with some severe hydrological hazards. These include glacial lake outburst floods, as well as damming of rivers by landslides resulting in large dam-burst floods (Hewitt 1982). The biggest anthropogenic influence on hydrology is Tarbela Dam at the outlet of the UIB. The dam is used for hydropower, supplying 20% of national electricity demand, as well as irrigated agriculture through the Indus Basin Irrigation System, which is the largest irrigation network in the world (Archer et al. 2010). With high rates of glacial and fluvial

erosion and transport in the steep valleys of the UIB, rapid reservoir sedimentation is a major issue at Tarbela Dam.

This brief introduction to the UIB thus demonstrates that the basin is characterised by its high mean and range of elevation, as well as its highly significant glaciated area and seasonal snow cover extent. The next sections of this chapter explain in detail how the climate, cryosphere and hydrology of the basin are related, according to our best understanding from the literature.

2.3 Climate

2.3.1 Precipitation

Much of the precipitation in the UIB occurs in winter and spring as snowfall, often originating from western disturbances (Archer and Fowler 2004; Fowler and Archer 2005; Palazzi et al. 2013). Western disturbances are essentially extratropical cyclones that can bring very heavy precipitation to Pakistan and northern India (Dimri et al. 2015; Hunt et al. 2018b). The subtropical westerly jet guides these synoptic scale low pressure systems eastward towards the UIB from their origins in the Atlantic and Mediterranean. As such, winter precipitation in the UIB has been shown to be well correlated with the North Atlantic Oscillation (NAO) index (Archer and Fowler 2004; Syed et al. 2006, 2010; Yadav et al. 2009). Filippi et al. (2014) found that the stronger jet during positive NAO phases intensifies western disturbances and moisture transport towards the Hindu Kush, Karakoram and western Himalaya (HKKH) ranges. Enhanced evaporation, particularly from the Arabian Sea, may be drawn in to provide a crucial source of moisture in some cases, thereby intensifying precipitation (Filippi et al. 2014; Tiwari et al. 2017; Hunt et al. 2018a).

Subtropical westerly jet intensity/position, western disturbance activity and precipitation in the western HKKH in winter also correlate with several other major modes of variability. These include the El Niño Southern Oscillation (ENSO) (Fowler and Archer 2005; Syed et al. 2006, 2010; Yadav et al. 2009, 2010; Cannon et al. 2017a), the Madden-Julian Oscillation (MJO) (Barlow et al. 2005; Cannon et al. 2017a) and the Polar Eurasia Pattern (Lang and Barros 2004). As Cannon et al. (2015, 2017a) note, interactions between modes influence both low level atmospheric states/fluxes and upper level circulation, which leads to substantial variation in the thermodynamic and dynamic conditions that shape the evolution of individual western disturbances. The underlying patterns driving temporal variability in winter and spring precipitation in the UIB are thus somewhat complex.

When winter and spring western disturbances reach the HKKH, its extreme elevation poses a substantial barrier. Orographic precipitation results, the intensity and duration of which is strongly influenced by cross-barrier wind speeds and moisture fluxes (Lang and Barros 2004). Variation in seasonal accumulations depends critically on whether the cyclones become terrain-locked in the notch formed where the Hindu Kush and Himalayan ranges meet, as this affects the time spent over the UIB and its southern slopes (Lang and Barros 2004; Norris et al. 2015; Dimri et al. 2015). Systems that do not become terrain-locked may pass more easily to the south and west of the western HKKH barrier, bringing precipitation farther east into the central Himalaya. Variations in the position/intensity of the subtropical westerly jet in relation to topography are strongly implicated in whether such terrain-locking occurs, which ultimately leaves a signature in precipitation distributions across the broader region (Schiemann et al. 2009). The hydrological implications of such variation can be enormous in the UIB, as evidenced by the drought years of the early 2000s (Archer et al. 2010).

Secondary to the winter/spring maximum, a lower precipitation peak is evident in summer at most observation locations inside the UIB, which are admittedly few in number (Archer and Fowler 2004; Fowler and Archer 2005). While some summer precipitation in the UIB is thought to be associated with occasional incursions of South Asian monsoon offshoots, westerly tracking depressions can also bring precipitation at this time of year, even though the subtropical westerly jet shifts to the north of the Tibetan Plateau (Fowler and Archer 2005; Hewitt 2014). The role of westerly activity in summer is confirmed by the correlation between the NAO and observed precipitation at this time of year, which led Archer and Fowler (2004) to recognise the importance of interplay between westerly and monsoonal circulation systems (Krishnan et al. 2009). Much is still unknown about the nature of this interaction, but recent work has shown how mid-tropospheric north-westerlies affect the advance of monsoon onset by forming a wedge of dry air that initially suppresses convection (Parker et al. 2016).

Interestingly, there is some suggestion from snow pits high in the accumulation area of the Biafo glacier that the proportion of annual precipitation falling in summer may actually be larger at high elevations than in valleys, where most measurements are taken (Wake 1989; Hewitt 2014). However, this is based on data from just one site for a short period, so substantial uncertainty remains regarding spatial and vertical variation in precipitation seasonality. More work has focused on spatial variation in mean annual precipitation. Winiger et al. (2005) used comparatively dense station networks in relatively small catchments to produce a logarithmic vertical profile. This increases from around 200 mm/a in valleys

(~2000 mASL) to more than 1500 mm/a at approximately 5000 mASL. The profile has formed the basis for a number of subsequent studies of UIB hydroclimatology (e.g. Bocchiola et al. 2011; Soncini et al. 2014; Reggiani et al. 2016). Drawing on the accumulation estimates from snow pits (Wake 1989), Hewitt (2014) show a similar albeit more linear profile, with precipitation increasing up to around 1000-2000 mm/a at 5000-6000 mASL before potentially decreasing at elevations above this due to exhaustion of moisture.

While extremely useful, these mean annual profiles do not capture the notable inter-annual variability in total accumulations, spatial variation in orographic gradients and the elevation of maximum precipitation, or general variations in precipitation magnitude across the large area of the UIB (Hewitt 2014). Indeed, the interplay of precipitation sources and topographic barrier effects has been observed to incur substantial horizontal gradients in precipitation at multiple scales (Young and Hewitt 1990; Reggiani and Rientjes 2015). Given the mismatch between scales of variation and observation network density, these gradients are very difficult to characterise fully using measurements, although correlations between stations in different parts of the UIB suggest the spatial variation to be fairly systematic (Archer and Fowler 2004). In an attempt to overcome the problem of sparse measurements, Dahri et al. (2016, 2018) augment precipitation observations with inferences and reports from glaciological field expeditions. Through kriging with external drift, Dahri et al. obtain a plausible spatial distribution of precipitation and reasonable consistency with runoff variation between observed sub-basins. However, notable uncertainties remain due to omission of evapotranspiration and sublimation terms, as well as approximations of the contribution made by net glacial mass balance.

In a contrasting approach, Immerzeel et al. (2012b, 2015) used coupled cryospheric-hydrological modelling to work out the vertical precipitation gradients and profiles required to reproduce glacier mass balance estimates from remote sensing. This is an ambitious method, subject to simplifications of key ablation and mass redistribution processes in the model, which cannot all be independently validated. Despite the large resulting uncertainty, as evidenced by error bar magnitudes of over a third of the estimates, the approximated precipitation amounts and distribution are more credible than many of the coarse gridded data products that are sometimes applied in the UIB. For example, catchment-scale precipitation amounts in the raw APHRODITE data product (Yatagai et al. 2012) were shown to be too low to give a water balance consistent with both estimated glacial mass balance and observed runoff (Immerzeel et al. 2015). This is intuitive, given that APHRODITE is essentially based

on interpolation of station data, which are skewed towards the drier locations of valleys. The Tropical Rainfall Measuring Mission (TRMM) products (Huffman et al. 2007) are also known to exhibit a substantial low bias for precipitation in the UIB, largely due to underestimation of snowfall (Forsythe et al. 2012a; Reggiani and Rientjes 2015; Maussion et al. 2014). In contrast, more reasonable basin-scale seasonality and amounts are often apparent in global reanalyses (Reggiani et al. 2017), but their resolution tends to be too coarse for direct hydrological applications.

From observed runoff and water balance considerations, Reggiani and Rientjes (2015) infer mean annual precipitation for the UIB at Tarbela Dam to be 675 ± 100 mm/a. Dahri et al. (2018) obtain an estimate of 541 mm/a, while Immerzeel et al. (2015) appear to obtain a median estimate of around 640 mm/a (Reggiani and Rientjes 2015). In conjunction with the relatively stable glacier mass balance in the basin (Section 2.4), the estimates by Dahri et al. (2018) may thus still be on the low side, once a reasonable amount of evapotranspiration is allowed for (Section 2.5.1). This highlights the substantial and persistent uncertainties regarding the amounts and spatiotemporal distribution of precipitation in the UIB.

While no estimates of inter-annual variation or trends in basin-scale precipitation are available, analyses have been conducted based on station observations. Archer and Fowler (2004) found no trend for stations with long records covering the whole of the twentieth century, but identified significant increasing trends in annual, winter and summer precipitation at some stations between 1961 and 1999. With the data record updated into the twenty-first century, Hasson et al. (2017) also found winter precipitation to be rising at both valley and high elevation automatic weather stations (AWSs), although the latter have shorter records (1995-2012). Interestingly, Hasson et al. also suggested spring precipitation to be decreasing, while noting that summer precipitation may have exhibited a shift from long-term wetting since the early 1960s to drying since the mid-1990s. However, as only 18 years of data are available since the mid-1990s, it is unclear whether these recent tendencies are an artefact of natural variability or the origins of longer term shifts.

The winter increasing trend could be consistent with studies identifying apparent rises in the frequency and strength of western disturbances, with a resultant upswing in extreme precipitation events (Cannon et al. 2015; Madhura et al. 2015). Interestingly, Yadav et al. (2009, 2010) found that the influences of the NAO and ENSO on winter precipitation variability in north-west India appear to have decreased and increased, respectively, in recent

decades. Yadav et al. proposed mechanisms by which western disturbances can be intensified under the warm phase of ENSO by passing through a cyclonic circulation anomaly in upper levels extending south-east from the Caspian Sea. The reduced influence of the NAO in winter is also borne out by the long records for the north-west Himalaya available to Bhutiyani et al. (2010), who suggest that the ENSO influence in summer has also weakened.

Additional trends analyses have highlighted the notable inter-station variation in trends. As only some locations are classified as exhibiting statistically significant trends, the coherence of the regional signal is difficult to assess (Khattak et al. 2011; Bocchiola and Diolaiuti 2013). Part of the issue may be sensitivities to the differences in analysis windows used by different authors, the importance of which was demonstrated by Forsythe et al. (2012b). Moreover, methodologically, the precipitation trend analyses conducted previously have been relatively simple. The possibility of change points in the series has not been explored, as well as the potential for complex series that involve both change points and trends (e.g. Guerreiro et al. 2014). The analyses may also be partly affected by fundamental criticisms of null hypothesis trend tests in hydrology, including flaws in methods commonly used to account for serial correlation, which have been raised recently by Serinaldi et al. (2018).

2.3.2 Temperature

Characterisation of the spatial and temporal variability in temperature in the UIB is paramount for hydrological modelling and water resources forecasting applications. Initial steps in this were taken by Archer (2004), who demonstrated that there is high spatial and vertical correlation of observed near-surface air temperatures on monthly, seasonal and annual time scales. Strong correlation was found to exist in fact for stations separated by large distances, suggesting that valley observations could provide a practically useful index of energy inputs across broad areas and vertical ranges, where finer scale variability is not measured. In trying to provide a more detailed picture of spatial variation, Mukhopadhyay and Khan (2016) attempted to quantify lapse rates by using single pairs of observation stations in gauged sub-basins of the UIB. They documented notable intra-annual (monthly) and spatial variation between catchments, but it is possible that relying on just two stations in each sub-basin is overly sensitive to local variability, for example due to aspect effects (Daly et al. 2008). This was recognised by de Scally (1997) using observations in the proximate Kunhar catchment. Differences in lapse rates below and above the snowline were shown to be crucial.

Similar to findings from elsewhere in the world (e.g. Pepin et al. 1999; Blandford et al. 2008), the importance and extent of diurnal, synoptic, seasonal and inter-annual variability in lapse rates have been measured in the well-instrumented Langtang catchment in Nepal (Immerzeel et al. 2014; Heynen et al. 2016). The challenges in making similar characterisations from the sparse observation network in the UIB may be partly addressed using alternative data sources, particularly remote sensing. Forsythe et al. (2012a,b) demonstrated that land surface temperature (LST) retrieved from the Moderate Resolution Imaging Spectroradiometer (MODIS) is a proxy for energy inputs in the basin. The relatively high spatial resolution of some MODIS LST data products could also help to better characterise variations related to topography (e.g. aspect, slope), as well as hydrologically critical quantities like the freezing isotherm. That said, the relationship between LST and air temperature varies depending on the time of day, season, land surface and synoptic conditions, such that considerable care is required in interpretation (Pepin et al. 2016). Estimation of air temperature from LST is a developing endeavour (e.g. Benali et al. 2012; Wenbin et al. 2013; Zhang et al. 2018), although ground data are still required for statistical model development and validation. If multi-decadal products can ultimately be developed by combining MODIS with longer records from the Advanced Very High Resolution Radiometer (AVHRR) instrument then characterisation of trends could potentially also be undertaken (Forsythe et al. 2012a,b).

To date, however, trend analysis has been based primarily on station observations, which show a number of interesting features relative to global trends. Fowler and Archer (2005, 2006) found that measured winter temperatures (daily means and maxima) have increased since the middle of the twentieth century, while summer temperatures (daily means and minima) have in fact decreased over the same period. This has resulted in a widening diurnal temperature range (DTR) in all seasons, in contrast to climate model projections and global trends. Forsythe et al. (2012b) demonstrated the continuation of these trends into early in the twenty-first century. Other studies have independently confirmed these findings (Khattak et al. 2011; Bocchiola and Diolaiuti 2013; Hasson et al. 2017). In addition, Hasson et al. (2017) suggested that a summer cooling tendency could also be identified since the mid-1990s in high elevation automatic weather stations, which were not used in Fowler and Archer's (2005, 2006) original work. The trend is also evident in global reanalyses, including ERA-40 (Quincey et al. 2009; Sharif et al. 2013) and ERA-Interim (Forsythe et al. 2015). The origins and likelihood of persistence of this summer cooling are major questions, as the implications for the basin's cryosphere and hydrology are potentially huge. Potential drivers are explored in Section 2.6.

2.3.3 Surface Mass and Energy Balance Controls

Studies of the full surface mass and energy balances in the UIB are limited due to poor data availability. Two studies attempt to use observed incoming shortwave radiation from the Water and Power Development Authority (WAPDA) automatic weather stations (AWSs) (Forsythe et al. 2015; Shrestha et al. 2015). However, the time series appear to exhibit severe and systematic under-estimation for a large part of the year, based on comparison with independent observations (Boerst 2013) and reanalyses (Forsythe et al. 2015). As such, seeking an alternative to local observations, Forsythe et al. (2015) used remotely sensed cloud cover fraction from MODIS with global reanalysis data to examine the modulating influence of cloud cover. This work characterised the seasonal, spatial and vertical variation in cloud cover, as well as the seasonally and diurnally varying connections between local cloud conditions and variability in near-surface air temperatures through cloud radiative effects. The possibility of diurnally asymmetric trends in cloud cover was also proposed as a contributor to the widening diurnal temperature range observed in the UIB.

While Forsythe et al. (2015) noted that an extension of the remote sensing record would be required to assess this more conclusively, potentially through a combination of MODIS and AVHRR, two recent studies have gained access to cloud cover observed at Pakistan Meteorological Department (PMD) weather stations (Bashir et al. 2017; Waqas and Athar 2018). Both studies report an increase in cloud cover in the summer months, which they relate to the observed summer cooling and widening diurnal temperature range through cloud reflection of incoming shortwave radiation. In principle this could be consistent with the explanations proposed by Forsythe et al. (2015), but the diurnal asymmetry in temperature (or cloud cover) change rates is not fully evaluated in either study. Moreover, as Forsythe et al. note, substantial complications are introduced by the correlation between daytime and night-time cloud cover, and the potential for daytime anomalies to "carry-over" to dominate the night-time signal in summer. This complexity results in a noisy relationship, such that it is hard to delineate the precise role of cloud cover at present.

The study by Bashir et al. (2017) also represents one of the only published evaluations of near-surface humidity and wind speed observations in the UIB. The authors show increases in both daytime and night-time humidity over recent decades, whereas wind speed observations show a decrease. Climatologies and inter-station variation are not reported. Other studies using available data on incoming radiation components, near-surface humidity and wind

speed are fairly few in number, generally focusing on glaciological applications or localised model input or evaluation, without climatological analysis (e.g. Mihalcea et al. 2006, 2008; Collier et al. 2013). There is thus substantial uncertainty in the spatiotemporal climatologies and variability of the climatic variables influencing (and being influenced by) surface mass and energy balances in the UIB. This is a clear problem for process-based evaluations of ablation, as well as wind transport and redistribution of snow.

2.3.4 Ways Forward

The previous sections have illustrated that substantial progress in understanding the climate of the UIB has been possible by using local observations, global reanalyses and remote sensing. However, it has also been argued that the spatiotemporal variability of hydrologically critical climate variables is still not well understood at a range of scales. Driven primarily by data paucity and limitations, as well as the restricted domains of applicability of complementary datasets, it continues to be difficult to describe climate fields at the relatively fine scales required in many types of process-based cryospheric and hydrological models, much like in most mountain regions in the world (Viviroli et al. 2011).

One possible response to this problem harnesses modern computing facilities to simulate near-surface climate using high resolution numerical weather models (NWPs) or regional climate models. These models simulate the range of climate variables needed for process-based cryospheric and hydrological studies, while maintaining physical consistency between variables. In various mountainous contexts, it has been shown that models like the Weather Research and Forecasting (WRF) model (Skamarock et al. 2008) can exhibit strong performance in hydrologically oriented applications, despite the complexity of terrain and variability of climate processes (Ikeda et al. 2010; Rasmussen et al. 2011; Prein et al. 2012; Silverman et al. 2013; Duethmann et al. 2013). The small but growing number of applications in the Hindu Kush, Karakoram and Himalayan ranges show similar potential (e.g. Collier et al. 2013, 2015; Norris et al. 2016). However, application and evaluation of this approach in the UIB is very limited, particularly at the basin and sub-basin scales of hydrological models. As such, we need to assess how well this relatively high resolution regional climate modelling can augment other datasets and improve our spatiotemporal climatologies in the basin. The High Asia Refined Analysis (HAR) is one such data product (Maussion et al. 2014), which has not been evaluated or applied in the UIB. This is introduced and discussed fully in Chapter 3.

2.4 Cryosphere

2.4.1 Glaciers

The most heavily glaciated part of the UIB is its share of the Karakoram mountain range (Figure 2.1). Glaciers here show a number of relatively unique features of relevance to hydrological and water resources studies. Firstly, avalanching on the extensive steep rockwalls is a key source of nourishment for a large number of glaciers, with avalanches often travelling well into the ablation zone (Hewitt 2011, 2014). From decades of field observation, Hewitt suggested that Karakoram glaciers may be classified by the relative contributions of avalanching, snowfall and wind-blown snow to accumulation, with some large glaciers in fact showing very limited accumulation areas compared with many alpine contexts. This largely prohibits application of basic glaciological measures, such as equilibrium line altitude (ELA) and accumulation area ratio (AAR), which are nevertheless still used in some studies (Mukhopadhyay and Khan 2016). Hewitt also suggested that the accumulation regime could be considered year-around, owing to summer storms and avalanching, while the ablation season is confined to summer. Ablation rates may be large due to high incoming shortwave radiation at this latitude, but sensitivity to summer weather is substantial. Storms can lead to sharp rises in albedo after snowfall, as well as suppressions in melt rates because of cloudiness. Capturing the influences of both snow redistribution and summer weather on accumulation and ablation would thus seem to be critical in water resources applications.

The extensive debris cover in the lower parts of Karakoram glaciers has also been a subject of substantial interest. Quantification of the effects of surface topographic variation and debris thickness on ablation in the field has been undertaken, which confirmed the melt-enhancing effect of thin debris cover and the melt-inhibiting role of thicker cover (Mihalcea et al. 2006, 2008). However, based on hypsometric considerations, Hewitt (2005, 2014) questioned whether the importance of debris cover has been overstated for water resources applications. Hewitt noted that, in general, most glacier ice (around 60 to 80%) is found between 3800 and 5800 mASL, with 3800-4800 being the critical elevation band for most meltwater generation. This zone is mainly comprised of thin debris cover, dusty ice or bare ice. Several investigations in other areas have shown the significance of supraglacial ponds and ice cliffs as hotspots for intense melting (Reid and Brock 2014; Pellicciotti et al. 2015; Miles et al. 2016), although the catchment-scale importance of these features is not well quantified in the

Karakoram. How to incorporate this partial understanding in hydrological studies is an open question.

These points begin to highlight the huge importance of the vertical dimension for understanding Karakoram glaciers and their relationships with changing climate conditions over massive elevation ranges (Hewitt 2011, 2014). These large elevation ranges give rise to complex and varied thermal, hydrological and dynamic glacial regimes. Measurements are limited, but there is thought to be a mixture of cold, temperate and polythermal conditions present, while block flow induced by basal sliding is likely to underpin a large part of glacier motion (Hewitt 2011, 2014). Some limited quantification of glacier flow rates has been undertaken from field observations and remote sensing (e.g. Smiraglia et al. 2006; Quincey et al. 2009; Scherler and Strecker 2012), while a complex mixture of hydrological and thermal drivers appear to induce surging in numerous different glaciers at different times (Quincey et al. 2011, 2015). However, the inaccessibility of many glaciers is a substantial challenge for developing an integrated understanding of the multiple processes and nuances that connect glacier mass balance and dynamics.

One of the most distinctive features of glaciers in the region is the so-called "Karakoram Anomaly" (Hewitt 2005). While other parts of the HKKH and Tibetan Plateau are generally following global trends of glacier recession, Karakoram glaciers appear to have shown relative stability overall in recent decades. Originally identified from observations of glacier expansions in the late 1990s (Hewitt 2005), several studies subsequently investigated and largely corroborated this finding using remote sensing for the early part of the twenty-first century (Bolch et al. 2012; Gardelle et al. 2013; Kääb et al. 2012; Gardner et al. 2013). These investigations variously used satellite gravimetry, laser altimetry and digital elevation model (DEM) differencing, but only covered a small number of years. Initial attempts to investigate the multi-decadal picture looked at area changes from satellite imagery (Bhambri et al. 2013; Rankl et al. 2014), which indicated the possibility of a shift from negative to stable or positive mass balance in the 1980s or 1990s, in broad agreement with observations (Hewitt 2005). More recently, digital elevation models have been constructed for multiple time slices and differenced to investigate mass changes since the 1970s. This has suggested that the Karakoram glaciers have largely shown stable or slightly negative mass balance for several decades (Bolch et al. 2017; Zhou et al. 2017).

While the anomaly was identified in the Karakoram, recent consensus is that it extends farther east (Cogley 2016; Azam et al. 2018). Based on ICESat laser altimetry for 2003-2008, Kääb et al. (2015) suggested that the anomaly of balanced/positive glacier mass balance is in fact centred on the western Kunlun Shan, with the Karakoram and Pamir mountains situated at the western margin of the anomaly. The findings from this short period study have been confirmed by repeat DEM differencing for a slightly longer time frame (2000-2016) (Brun et al. 2017). Wang et al. (2018) recently identified that area and surface elevation changes in the western Kunlun Shan since the 1970s are negligible. In conjunction with results for the same period from the Karakoram (Bolch et al. 2017; Zhou et al. 2017), this suggests that the glacier mass balance anomaly shows a relatively persistent spatial footprint over several decades. This raises some interesting questions about potential drivers of the anomaly, as discussed in Section 2.6

2.4.2 Snow

Snow processes in the UIB play critical roles with respect to glacier behaviour, hydrology and land-atmosphere interactions. The importance of snow redistribution by avalanches and wind for glacier accumulation has already been emphasised in Section 2.4.1 (Hewitt 2011, 2014), while Section 2.5 will show the huge contribution of seasonal snowmelt to river flows (Archer 2003; Mukhopadhyay and Khan 2015). As such, some attempts to investigate certain processes were undertaken several decades ago in the upper Kunhar catchment situated in the Jhelum basin, a tributary of the lower Indus that drains the southern slopes of the Himalaya. Particular emphasis was placed on quantifying avalanche mass (de Scally and Gardner 1989), the relative ablation rates of avalanched and undisturbed snow (de Scally and Gardner 1990), and the effects of avalanching on snowmelt runoff at the catchment scale (de Scally 1992).

Ablation rates were found to be generally higher than in other mountain regions (de Scally and Gardner 1989). Reasonably high degree day factors for snow of around 6 mm/°C/day were also measured in the Garhwal Himalaya, to the south-east of the UIB, but the values are not necessarily outside of ranges reported from other regions (Singh and Kumar 1996; Singh et al. 2000) or indeed from other parts of the Himalaya (Hock 2003). Avalanche deposits showed higher surface melt rates compared with undisturbed snow, but the net effect of avalanching was in fact to delay runoff, as the deposits were concentrated in small areas of deeper, persistent snow (de Scally and Gardner 1989). For the Kunhar basin, it was estimated

that 6% of total runoff could be from avalanched snow following heavy winters, with 2 to 3 months delay in snow disappearance possible relative to undisturbed snow (de Scally 1992).

The work also confirmed that air temperature was strongly correlated with snowmelt, but the scatter in the relationship was large enough to give quite substantial error bounds for basin-scale predictions, especially in the case of undisturbed snow, which was much more areally extensive (de Scally and Gardner 1989; de Scally 1992). Interestingly, flux measurements suggested a role for condensation as a source of melt energy, due to high near-surface vapour pressures following monsoon onset. This is somewhat of a contrast to common conditions where sensible and latent heat fluxes largely cancel each other, but measurements of the full surface energy balance were unavailable to explore this issue further unfortunately. Given the uncertainties over monsoon penetration into the HKH, as well as very limited characterisation of spatiotemporal variation in humidity, it is unclear how significant this process could be in the UIB.

However, with the exception of a few studies like these ones, investigations of snow processes and snow dynamics in (or near) the UIB are limited, especially in comparison to glaciological studies (Hewitt 2014). This is in fact also true in much of the Himalayan arc, largely due to lack of observational data. A case for vastly expanding in-situ snow observations in the region was put forward by Rohrer et al. (2013), who demonstrated the need for local data to constrain remote sensing products in the face of potential changes in the surface energy balance resulting from climate change. In the absence of such observations, researchers have depended heavily on remote sensing. Forsythe et al. (2012a,b) characterised the annual cycle of snow-covered area (SCA) and its inter-annual variability using MODIS data products. They also provided some validation of the dataset using local observations and process-based reasoning, while showing the substantial potential of remote sensing data as a predictor in hydrological applications at multiple scales. Tahir et al. (2011a, 2015) also analysed MODIS SCA in relation to UIB hydrology, although their trend analysis was conducted for a very short record period. Trends in snow water equivalent (SWE) over a longer period were estimated using passive microwave remote sensing by Smith and Bookhagen (2018), but the absolute accuracy of these trends is likely to be limited by difficulties with SWE retrieval in complex terrain

While remote sensing data have led to important developments in our understanding of snow cover dynamics, lots of unanswered questions remain with respect to snow processes in the

UIB. The energy balance of the snowpack and its spatiotemporal variability are poorly quantified, along with the relative magnitudes of melt and sublimation. Studies in other parts of the Himalaya and other high mountain regions suggest that sublimation could be large, particularly when considered jointly with wind transport (Strasser et al. 2008; Macdonald et al. 2010; Ayala et al. 2017; Stigter et al. 2018). Meanwhile, limited analyses of the surface energy balance composition mean caution is required in inferences about responses to future conditions or selection of model types. The importance of high incoming shortwave radiation has been highlighted, but studies of albedo and its variation at multiple scales are largely absent. Uncertainties in patterns and magnitudes of snow redistribution by avalanches and wind are also large.

2.5 Hydrology

2.5.1 Catchment Understanding

Estimates of the mean annual water balance terms for the UIB at Tarbela Dam are given in Table 2.1. This suggests there to be reasonable agreement between the central estimates of both precipitation and evapotranspiration by Reggiani and Rientjes (2015) and Immerzeel et al. (2015). Estimates in the former study were derived primarily from water balance considerations, whereas estimates in the latter were based on inverse modelling to reconstruct precipitation from observed glacier mass balance and river flow (Section 2.3.1). While these estimates of precipitation and evapotranspiration are consistent with each other, uncertainty remains regarding their accuracy. Some degree of compensating error is likely, given that the precipitation and evapotranspiration terms cannot be validated independently. However, the estimates do represent an advance over earlier studies in which precipitation was severely underestimated, leading to implausible rates of glacier retreat (Reggiani and Rientjes 2015). Dahri et al.'s (2018) estimate of precipitation reported in Table 2.1 may still be a little low, but it is still more plausible than a precipitation rate of ~300 mm/a suggested in earlier investigations (e.g. Immerzeel et al. 2009, 2010).

Three primary hydrological regimes can be identified in the UIB: glacial, nival and pluvial (Archer 2003). The higher parts of the basin give rise to the glacially dominated river flows of the Hunza, Shigar and Shyok sub-basins. Elsewhere, runoff mainly originates from seasonal snowmelt, derived from snowfall in the preceding winter and spring, or rainfall in the concurrent season. The coherent spatial differentiation of these glacial, nival and pluvial regimes stems in large part from variations in hypsometry and interactions of topography with

westerly and monsoonal weather systems. In each regime, runoff is low during winter, restricted to a baseflow with little fluctuation at sub-seasonal scales. River flows begin to rise in spring and peak in summer in all cases, although the precise timing of the peak varies depending on the relative contribution of glacial, nival and pluvial sources in a given catchment. As evidenced by comparing the annual cycle of catchment area with night-time temperatures above freezing (continuous melt area, CMA) with that of river flow, runoff in nival regimes tends to be limited by mass available for ablation, while energy inputs are the limiting factor in glacial regimes (Archer 2003; Forsythe et al. 2012b). In the UIB this tends to lead to relatively earlier and later runoff peaks in nival and glacial catchments, respectively, while inter-annual variability in flow rates and volumes can be very substantial in both cases.

Source	Precipitation (mm/a)	Evapotranspiration (mm/a)	Runoff (mm/a)
Reggiani and Rientjes (2015)	675 ± 100	200 ± 100	460
Immerzeel et al. (2015)	640	180*	
Dahri et al. (2018)	541	81*	

Table 2.1 – Mean annual water balance terms for the Upper Indus Basin at Tarbela Dam from the recent literature. Precipitation and evapotranspiration are estimated, but runoff is observed (1961-2009 reference period). * indicates evapotranspiration calculated by subtracting observed runoff from estimated precipitation (where evapotranspiration was not reported in the study).

Some attempts have been made to quantify the relative contributions of different components of river flow using hydrograph separation. For example, Mukhopadhyay and Khan (2015) estimated that on average 70% of the total flow reaching Tarbela reservoir is meltwater, of which 26% originates from glacier melt and 44% from snowmelt. The method underlying these estimates divides a catchment into three elevation bands, with simplifying assumptions about the relative contributions from snow and glaciers in each band for each month of the year. Rainfall-runoff is implicitly lumped in with the substantial baseflow components reported. The approach, informed by conceptual understanding, usefully takes into account catchment hypsometry and the vertical distribution of glacial areas. As such, it may give reasonable, rough approximations, but the uncertainties are likely to be significant, if unquantified so far. However, one of the alternative approaches to estimating flow components, numerical modelling, also leads to uncertain and substantially varying estimates. Taking the high elevation Hunza sub-basin as an example, model estimates of the glacier melt

contribution to mean annual flows range from 33% (Shrestha et al. 2015) to an implausibly high 85% (Lutz et al. 2016a), compared with 43% from the aforementioned hydrograph separation technique (Mukhopadhyay and Khan 2015). This reflects the ongoing modelling challenges in the UIB discussed in below, as well as the persistent uncertainties regarding the constitution of UIB river flows.

While these uncertainties are important, seasonal river flows can still be forecast with a degree of skill sufficient to aid water managers. Fowler and Archer (2005, 2006) demonstrated that summer runoff in snow-dominated sub-basins is closely related to winter precipitation measured at valley stations. In contrast, in heavily glaciated sub-basins, it was found to be strongly associated with summer temperatures, again based on relatively low elevation observations. These findings support the idea that summer runoff in snow- and glacier-dominated catchments in the UIB is primarily limited by mass and energy constraints, respectively. They are also in agreement with the spatial correlation of observations, which show strong relationships in winter precipitation even for widely separated stations (Archer and Fowler 2004), as well as high correlations for summer temperatures at monthly and longer time scales (Archer 2004). Archer and Fowler (2008) used these findings to develop multiple linear regression models for seasonal river flow forecasting in the neighbouring Jhelum basin. Although the basin is influenced by the South Asian summer monsoon, winter (rather than summer) precipitation was found to be a key predictor of spring and summer runoff, in agreement with an earlier study (De Scally 1994).

The success of flow forecasting here suggests that, although sparse, the observation network in the UIB and neighbouring basins provides significant information content for hydrological applications. Forsythe et al. (2012a,b) extended this using MODIS remote sensing data products. They found that quantities derived from snow-covered area (SCA) and land surface temperature (LST) products provide adequate proxies for anomalies in precipitation and near-surface air temperature, respectively. As such, they also demonstrated the potential for these products to be used in hydrological forecasting at multiple scales.

Previous hydrological analyses have also examined trends in river flows. Fowler and Archer (2006) found that summer (and therefore annual) flows in the high elevation, heavily glaciated Hunza catchment draining the western Karakoram declined substantially since the mid-1960s when records began. In contrast, river flows in the Shyok catchment farther east in the Karakoram did not show such a clear reduction. Mukhopadhyay et al. (2015) reached

similar conclusions for the Hunza catchment with updated flow records. Conversely, for the Shyok sub-basin, Mukhopadhyay et al. suggested that summer flows may be increasing. However, they only found these trends to be statistically significant after smoothing the data, whereas trends in the Hunza catchment were identified as significant with both raw and smoothed data. For the Shigar basin, which flows down from the central Karakoram, Mukhopadhyay and Khan (2014) again inferred rising river flows from smoothed series alone.

With Mukhopadhyay and Khan's (2014) and Mukhopadhyay et al.'s (2015) analyses conducted on (stratified) monthly time series (i.e. only 30-40 data points in the context of substantial inter-annual variability), applying trend analysis to smoothed series may not be sufficient to form definitive conclusions. This is particularly the case for the Shigar basin, where nearly half of the flow series was estimated from another gauging station. While reported correlations between the gauges may be reasonably high (explaining between 35 and 77% of total variance in summer months), the (unquantified) errors are still likely to be large enough to significantly affect procedures as sensitive as trend analyses. More generally, approaches based on resampling or simulation may provide a better means of seeing whether any notable trend is present in the context of such variability.

Sharif et al. (2013) also noted that middle elevation nival catchments tended to show increasing summer and annual flows, although the statistical significance of trends was found to be quite variable. This study also showed that trends in runoff timing are not clear or consistent between catchments. The picture of hydrological change is thus complicated and raises the question of how much meaningful signal can be identified from the noisy time series. Indeed, Sharif et al. note that even where strong trends are apparent, such as in the Hunza catchment, inter-annual variability tends to far exceed trend magnitude. Therefore, historical hydrological variability would likely provide a reasonable basis for short- or medium-term water resources and flood risk planning.

This discussion suggests that we have a reasonable understanding of hydrological regimes, sources of river flows, and climatic controls on summer runoff in snow- and glacier-dominated catchments and their spatial correlation. This provides good practical means for seasonal river flow forecasting. Trends are ambiguous to a degree, with heavily glaciated catchments most likely showing declining or stable river flows and nival catchments potentially showing increasing flows in some cases. Runoff timing trends are messy. However, this means that numerous gaps in our knowledge are present. Firstly, catchment

water balances are still poorly quantified. Observed runoff data are available for relatively large sub-basins, but the spatiotemporal distribution of precipitation is still uncertain (Section 2.3.1) and evapotranspiration/sublimation estimates are few in number. Secondly, the relative contributions of different sources of river flow are still problematic, as evidenced by the large range of reported estimates in the literature. Thirdly, flow pathways and travel times are not well constrained. Data on surface and subsurface hydraulic properties are absent, with the role of substantial valley superficial deposits (glaciofluvial sediments) in the hydrological system little known. Finally, we need to understand the origins of the complex hydrological trends, in order to better understand how to manage water resources in the future. This is discussed in Section 2.6.

2.5.2 Modelling

In emphasising the need for uncertainty analysis in hydrological modelling, Hrachowitz and Clark (2017, p. 3966) recently suggested that "[w]e are currently in a position where we, in an exaggerated way, feed wrong models with wrong input data and calibrate them to wrong output data to obtain wrong parameters". This is of course a simplification, and one which should not undermine the invaluable role of modelling in complex terrain and elsewhere (Burlando et al. 2002). However, it does allude to the challenges present in each aspect of modelling basins like the UIB. Amplified in mountain regions (Klemeš 1990), the often substantial uncertainties in input and evaluation data mean that in most cases we can only partially characterise the array of interacting processes and spatiotemporal variability in a catchment. In conjunction with the simplifications of reality necessary in models, it can thus be difficult to know if a given model is the most appropriate one, if it produces the right results for the right reasons, or if it is truly justified by available data. This has led to longstanding and unresolved debates on the appropriateness of different model formulations for different applications across the field of hydrology (e.g. Beven 1989; Grayson 1992; Beven 2001, 2002, 2006; Wood et al. 2011; Beven and Cloke 2012; Montanari and Koutsoyiannis 2012; Fatichi et al. 2016; Savenije and Hrachowitz 2017; Clark et al. 2017; Hrachowitz and Clark 2017).

For the UIB specifically, we contend that its particularly severe data paucity lies at the heart of barriers to modelling improvements. For example, similar to other mountain contexts, the acute limitations in observation coverage, accuracy and representativeness mean that climate inputs to hydrological models need to be modelled themselves (Klemeš 1990). Most prior

modelling studies in the UIB and broader Himalayan region have taken simple approaches to this challenge. Temperature fields are typically distributed in space based on elevation alone, using lapse rates and time series from one or more observation locations. A similar approach is often taken with precipitation, with vertical gradients sometimes approximated from published orographic profiles (e.g. Bocchiola et al. 2011) and sometimes estimated through calibration procedures (e.g. Ragettli et al. 2013). Troublingly, data products with known limitations in reproducing precipitation amounts and gradients have also been used in some cases (see e.g. Immerzeel et al. 2009, 2010; Reggiani and Rientjes 2015), although the need to simulate a credible water balance is now more widely understood (Immerzeel et al. 2015). Other studies have used remote sensing products, often with snow-covered area (SCA) as an input instead of precipitation (e.g. Tahir et al. 2011b), which has the disadvantages of not tracking the water balance or being easily applied in climate change impact studies. Finally, a small number of studies have attempted to use statistical downscaling of coarse reanalysis products (e.g. Brown et al. 2014). However, there are not many such studies in the Himalaya, and the single example from the UIB appears to have used erroneous incoming radiation data in its downscaling (Shrestha et al. 2015).

Most of these observation-based approaches are restricted to more commonly collected climate variables, namely precipitation and temperature. To some extent this limits the type of models that can then be applied. This fits with the tendency for applying simple modelling approaches in the UIB identified in Chapter 1. Focusing on snow and ice ablation, the primary source of river flows in the UIB, the algorithms applied in the majority of models are typically versions of temperature index (TI) (e.g. Bocchiola et al. 2011; Rees and Collins 2006; Immerzeel et al. 2009, 2012a, 2013; Soncini et al. 2014) or enhanced temperature index (ETI) approaches (e.g. Ragettli et al. 2013, 2015; Immerzeel et al. 2014; Minora et al. 2015). The former scales melt with air temperature following exceedance of a critical temperature threshold, while the latter additionally incorporates a term for net shortwave radiation at the surface. Many of the aforementioned studies consider these approaches to perform satisfactorily with respect to observed river flows, remote sensing of snow cover dynamics and, increasingly, regional glacier mass balance estimates from geodetic or other methods.

Yet, as essentially semi-empirical abstractions from the surface energy balance, there are several unresolved questions relating to these methods. This includes the extent to which the ETI method provides better performance or robustness compared with the classical TI method, especially in different climatic settings (e.g. MacDougall et al. 2011; Vincent and Six

2013; Gabbi et al. 2014; Réveillet et al. 2017). For both methods, the stability of parameters in space and time can also be questioned (Matthews et al. 2015), while sublimation and its enhancement by wind transport is usually neglected.

The corollary to simple modelling approaches being preferred in the UIB is that more complex approaches are little explored. These include methods predicated on solution of the surface energy balance. Only one such attempt has been made in the UIB (Shrestha et al. 2015), albeit with apparently erroneous incoming radiation inputs. Applications elsewhere in the Himalayan region are mainly restricted to short point-scale simulations driven with AWS measurements to help calibrate ETI parameters (Ragettli and Pellicciotti, 2012; Pellicciotti et al., 2012; Ragettli et al., 2013; Ragettli et al., 2015). Yet process-based energy balance models may confer advantages for testing hypotheses about processes, completing mass and energy balance frameworks in coupled land-atmosphere models, and making projections in non-stationary climates. These potential advantages arise because of a closer approximation of the physics thought to best describe the relevant processes. However, in line with broader debates in hydrology (Beven 2012), there are questions as to whether higher complexity may be reasonably borne out by the data available for model input, parameter assignment and evaluation. Indeed, based on model inter-comparisons in other regions, it is likely that only groups of generally well-performing process-based model configurations are identifiable (Etchevers et al. 2004; Rutter et al. 2009; Essery et al. 2009, 2013; Magnusson et al. 2015; Lafaysse et al. 2017).

The multiple climate variables needed in more complex models can be estimated from various parameterisations, but their accuracy may vary between contexts and their applicability is not always easy to determine in data sparse regions. Moreover, observation-based approaches tend to substantially simplify spatiotemporal variability where station density is low; the degree to which this is acceptable for a given application is often unclear. Elsewhere in the Himalaya, in small catchments where station density is higher, observation-based estimation of variability has been shown to give good performance, while permitting inferences about catchment behaviour (Ragettli et al. 2015). However, how to upscale to larger basins is an outstanding question, while most of the UIB does not have any detailed measurement networks in smaller scale catchments (at least where the data is in the public domain). These issues reinforce the need to investigate further the potential for high resolution regional climate modelling to augment other data sources in the UIB (Section 2.3.4)

The issues raised here are introduced in more detail in Chapters 3, 4 and 5, which provide additional discussion of the relevant literature. However, in concluding this section, we note that there is substantial scope for more model inter-comparison in the UIB and surrounding regions. There has been little work comparing simple and more complex model configurations. It is thus timely to consider whether recently developed data products enable the application of more process-based models, and whether this confers notable advantages in different applications.

2.6 Climate Change

As noted in Section 2.4.1, the "Karakoram anomaly" was initially designated as an essentially glaciological phenomenon. Specifically, the term referred to the relative stability of glacier mass balance compared with regional and global signals of glacier retreat. However, based on the climatic and hydrological data discussed above, it seems reasonable to assert that this region shows anomalous behaviour in other respects too. Climatically, this pertains particularly to the summer cooling temperatures and widening diurnal temperature range, driven particularly by declining daily minimum temperatures (Section 2.3.2). The significance of trends in observed precipitation is variable, although some increase may have occurred in winter. Meanwhile, depending on which part of the range is considered, river flows appear to have fallen or been roughly stable in the heavily glaciated catchments draining the Karakoram (Section 2.5). How these climatic, glaciological and hydrological patterns might be connected, as well as the nature of potential underlying drivers, are thus critical questions.

Different studies have emphasised different aspects of this multi-dimensional anomaly, while also offering different explanations of its causes. Initial attempts to understand the anomaly focused primarily on glaciological factors, including the thick debris cover that can be found on long, relatively low-lying glacial tongues. Based on remote sensing of terminus dynamics, Scherler et al. (2011) suggested that faster rates of glacier retreat were found in areas with less extensive debris cover, whereas thick debris cover in the Karakoram was associated with terminus stability. However, subsequent work showed that rates of surface lowering of debris-covered and clean ice glacier tongues were often in fact similar (Gardelle et al. 2012; Bolch et al. 2012). The factors contributing to this are thought to include "hotspots" of ablation on debris-covered glaciers, namely supraglacial ponds and ice cliffs, which partly compensate for the low melt rates under thick debris elsewhere, as well as lower fluxes of ice into slow-moving debris-covered ablation zones with shallower gradients (e.g. Brun et al. 2018).

Thus, despite its importance, an explanation of the anomaly based on debris cover alone appears incomplete so far, especially as it does not consider the possible relationships between glaciological, climatic and hydrological behaviour in the region. Hewitt (2005, 2011) suggested instead that verticality and the differentiation of myriad processes by elevation may be key to the Karakoram's anomalous behaviour, noting that only some of the processes and features (or some interaction of them) would be particular to the range. One such feature is the exceptional hypsometry of the Karakoram, with its unique fraction of very high elevation areas, which are additionally subject to a year-round accumulation regime. Summer snowfall would both nourish glaciers with additional mass and reduce ablation through increasing surface albedo and cloudiness. Fowler and Archer (2006) and Hewitt (2011) noted that the sensitivity of ablation to summer weather could mean that the observed trends of summer cooling contribute significantly to reduced ablation (see also Quincey et al. 2009), as manifest in observed river flows in some high elevation, heavily glaciated sub-basins (Section 2.5).

This explanation essentially puts climate as the principal driver of recent glacier and river flow behaviour in glacially dominated UIB regimes. Other studies have also foregrounded the role of climate, albeit with different emphases and often with reference to only some dimensions of the climatic, glaciological and hydrological anomaly. For example, Kapnick et al. (2014) used relatively high resolution modelling to suggest that the larger winter snowfall in the Karakoram helps to make it less sensitive to warming than the Himalaya, such that the difference in seasonal precipitation cycles is important. Ridley et al. (2013) also emphasise precipitation, finding that increases in winter western disturbance activity may maintain a comparatively high mass input to Karakoram glaciers. Other authors suggest that rising summer precipitation could be important in directly supporting runoff, albeit based only on trend analysis using reanalyses rather than observations (Mukhopadhyay et al. 2015). The co-variation of trends in cloudiness, humidity, precipitation and wind speed with falling summer temperatures has also recently been studied, albeit based on data from low elevations stations alone (Bashir et al. 2017). The authors hypothesise that increasing summer cloudiness has limited temperature rises and reduced ablation in the Karakoram. However, the drivers behind trends across multiple climate variables are not considered.

One attempt to provide an integrated explanation of the climatic, glaciological and hydrological anomaly is the Karakoram Vortex (KV). As introduced in Chapter 1, Forsythe et al. (2017) introduced the KV as a means of better understanding the regionally differentiated trends across the Himalaya and Tibetan Plateau. A fuller description is provided in Chapter 6,

but in essence the KV describes anomalous circulation patterns centred on or near the Karakoram throughout the year. These anomalies are related to variations in the position and intensity of the subtropical westerly jet, as well as the South Asian monsoon in summer. A key feature of the KV is that its circulation anomalies contract from winter to summer, resulting in a dipole in climate response between the Karakoram and the eastern Himalaya and Tibetan Plateau. In conjunction with trends in westerly and monsoonal circulation, this may be critical for understanding divergent climate and glacier trajectories. Similar findings were recently reached by Mölg et al. (2017) and Norris et al. (2018). These explanations usefully start to relate regional circulation to local climate patterns, recognising the importance of interplay between westerly and monsoonal systems (Krishnan et al. 2009; Bothe et al. 2011; Saeed et al. 2011; Wei et al. 2017). However, much is still to be investigated, including the drivers of regional circulation variability and trends, as well as the mechanisms leading to sub-regional differentiation in climate trajectories.

In terms of projections for the future, some general tendencies are common in most previous projections. These include initial increases in summer river flows due to higher rates of glacier melt in a warmer world. This is ultimately followed by a reduction in summer flows following glacier wastage, although this tends to be accompanied by a shift in the annual hydrograph involving higher flows in spring (e.g. Immerzeel et al. 2013; Lutz et al. 2014; Soncini et al. 2014; Ali et al. 2015). However, the rates and magnitudes of changes in the timing and magnitudes of river flows vary notably in model-based studies, particularly due to the wide range of projections from different climate models (e.g. Ragettli et al. 2013; Lutz et al. 2014, 2016a,b; Palazzi et al. 2015). Spread in both precipitation and temperature projections is very large (e.g. Palazzi et al. 2015; Lutz et al. 2016b), although most models tend to project warming. Uncertainties in potential climate and hydrological trajectories for the UIB are thus very large (Lutz et al. 2016a,b).

In order to understand and potentially constrain this spread, it is first critical to evaluate how well GCMs simulate key features of the regional climate, and whether there are connections between model skill, process representations and projections. Previous studies have evaluated a number of pertinent aspects of GCM behaviour over the Himalaya and Tibetan Plateau. These include: precipitation and temperature climatology (e.g. Su et al. 2013; Palazzi et al. 2015; Hasson 2016a; Hasson et al. 2016); land- and near-surface processes, feedbacks and elevation-dependent warming (e.g. Rangwala et al. 2013; Ghatak et al. 2014; Rangwala et al. 2016; Palazzi et al. 2017; Guo et al. 2018); monsoon dynamics (e.g. Sperber et al. 2013;

Levine et al. 2013; Saha et al. 2014; Wang et al. 2017); and other relevant facets of global and regional atmospheric circulation (e.g. Lee et al. 2014; Xu et al. 2017). However, substantial gaps remain. These include identifying whether climate models reproduce some of the aforementioned processes likely to lead to the unique behaviour of the Karakoram and UIB over recent decades. This is examined at length in Chapter 6.

While this study focuses particularly on KV regional circulation influences on climate, cryospheric and hydrological trajectories in the UIB, there are other factors that may also have important influences. With their pronounced dynamic and thermodynamic effects across large parts of Asia, aerosols are one such factor. Aerosols affect the flux of shortwave radiation reaching the surface, convective potential and atmospheric stability, atmospheric heating profiles, cloud microphysics and precipitation (Ramanathan et al. 2001, 2007; Li et al. 2016). This leads to effects on regional circulation patterns, including monsoon onset, development and evolution (Bollasina et al. 2013). Indeed, aerosols are thought to be implicated in the weakening of the South Asian summer monsoon circulation over recent decades, as well as the accompanying precipitation decrease over India (Bollasina et al. 2011, 2014; Li et al. 2016). Given the important interplay between the subtropical westerly jet and the monsoon in shaping UIB summer climate, as described earlier, regional patterns and trends in aerosols are therefore of relevance to climate change and water resources in the basin.

However, some of the more localised effects of aerosols on climate may be comparatively restricted in the UIB, owing to the lower aerosol optical depth in this part of Asia relative to regions farther to the south and east (Li et al. 2016). Nevertheless, there is still potential for some hydrologically significant aerosol influences, such as periodic deposition of black carbon or dust to reduce snow and ice surface albedo and increase melt rates (Gertler et al. 2016). There may be variation within the UIB with respect to the occurrence and relevance of this process, with deposition potentially more prevalent in the western Himalaya than the Karakoram (Gautam et al. 2013). While it is clearly important to consider possible aerosol and deposition effects on UIB climate and surface properties, the issues represent an additional layer of complexity that should be studied after progress has been made on the core research gaps guiding this work, as outlined in Chapter 1.

2.7 Conclusions

The review suggests that, despite vital progress in many areas, uncertainties in the spatiotemporal variability of surface and near-surface climate in the basin are still large. It is

contended that regional climate modelling at moderate or high resolution represents a valuable but under-utilised source of data to augment local observations, remote sensing and global reanalyses in this context. With the increasing potential for these datasets and methods to partly address the data gap, it is suggested that evaluations of the applicability of process-based models solving the full surface energy and mass balances are now needed. The ongoing difficulty in identifying appropriate model structures and process representations, even in well-instrumented contexts, suggests the need for systematic modelling frameworks to explore the influence of modeller decisions and identify where the ensemble of modelling possibilities can be constrained. The review also finds that there is increasing understanding of how the interplay and evolution of mid-latitude westerly and monsoonal circulation systems are shaping the unique climatic, cryospheric and hydrological trajectories in the basin. However, future trajectories remain highly uncertain. Further process-based evaluations of climate models are thus required to see if projections can be better constrained.

Chapter 3

Near-Surface Climate Representation in the High Asia Refined Analysis

3.1 Introduction

Runoff generated in the high mountains of the Upper Indus Basin (UIB) is a vital resource for vast populations. Combined with concerns about the impacts of climate change, this has led to substantial interest in the UIB's climate and hydrology over many decades (e.g. Young and Hewitt 1990; Archer et al. 2010). Yet here, like most mountain regions in the world, data paucity remains a persistent and major challenge for researchers and practitioners (Viviroli et al. 2011). Measurement networks are sparse and focused on lower elevation valley locations, while data quality and continuity issues are inherent in such a harsh and remote environment. This hinders application of one of the key tools for research and management in the basin, namely process-based cryospheric and hydrological modelling.

One response to this problem harnesses modern computing facilities to simulate near-surface climate in the Himalaya using high resolution numerical weather models (NWPs) or regional climate models. For example, Collier and Immerzeel (2015) found that the Weather Research and Forecasting (WRF) model (Skamarock et al. 2008), with a 1 km resolution inner domain, reproduces the key features of variability in a detailed measurement network in the Langtang catchment, Nepal. Similarly, Karki et al. (2017) demonstrate the added value of convection-permitting simulations elsewhere in the Nepalese Himalaya. Good correspondence between WRF and observations was also found by Collier et al. (2013, 2015) in simulating glacier-atmosphere interactions in the Karakoram, and by Norris et al. (2016) in their study of the spatiotemporal distribution of precipitation across the Himalayan arc.

The relatively small number of WRF simulations conducted to date for this region therefore show clear potential to supplement local observations. While most of these experiments are for reasonably short (<1 year) simulation periods, Maussion et al. (2014) performed high resolution WRF simulations of 14 years for the whole of the Tibetan Plateau and its adjoining mountain ranges. The resulting publicly available dataset is the High Asia Refined Analysis (HAR), which is in essence a dynamical downscaling of coarser global analysis using 30 and 10 km nested domains. For this region, the HAR thus represents a uniquely fine spatial discretisation for a WRF simulation of this length over such a large area. As such, it has

recently been used to improve our understanding of spatial and temporal patterns of precipitation variability and atmospheric water transport (Maussion et al. 2014; Curio et al. 2015; Curio and Scherer 2016). The HAR has also been employed to successfully provide offline meteorological forcing data for models of glacier mass balance and hydrology (Huintjes et al. 2015; Biskop et al. 2016; Tarasova et al. 2016), as well as to examine the connection between mid-latitude westerlies and glacier mass balance in monsoonal parts of the region (Mölg et al. 2014).

In their evaluation of HAR precipitation skill, Maussion et al. (2014) found that the dataset corresponds well with local observations, certain features of the Tropical Rainfall Measuring Mission (TRMM) product (Huffman et al. 2007), and known relationships between precipitation and topography. Our aim in this paper is to extend the evaluation of the HAR dataset in several ways. Firstly, we focus in detail on the near-surface climate of the UIB, an area for which ground-based measurements were unavailable for Maussion et al.'s (2014) original assessment. Secondly, we go beyond precipitation to consider other climate variables related to surface mass and energy balances, including temperature, humidity and incoming radiation. Finally, we explore some of the key factors affecting HAR performance. For this, we concentrate on cloud and albedo influences, based on the nature of bias patterns identified in near-surface variables.

The primary focus in our evaluation is on HAR climatological representations and biases, as reasonable performance in these respects is vital for numerous applications. The findings from this analysis can therefore inform future hydrological and cryospheric modelling and other applications in this part of the Himalayan arc, as well as regional WRF modelling.

3.2 Study Area

The UIB, shown in Figure 3.1, exhibits pronounced hydroclimatic variation on a range of scales. The higher parts of the basin give rise to the glacially dominated hydrological regimes of the Hunza, Shigar and Shyok sub-basins (Sharif et al. 2013). Elsewhere, runoff mainly originates from seasonal snowmelt, derived from snowfall in the preceding winter and spring, or rainfall in the concurrent season (Archer 2003). The coherent spatial differentiation of these glacial, nival and pluvial regimes stems in large part from variations in hypsometry and interactions of topography with westerly and monsoonal weather systems.

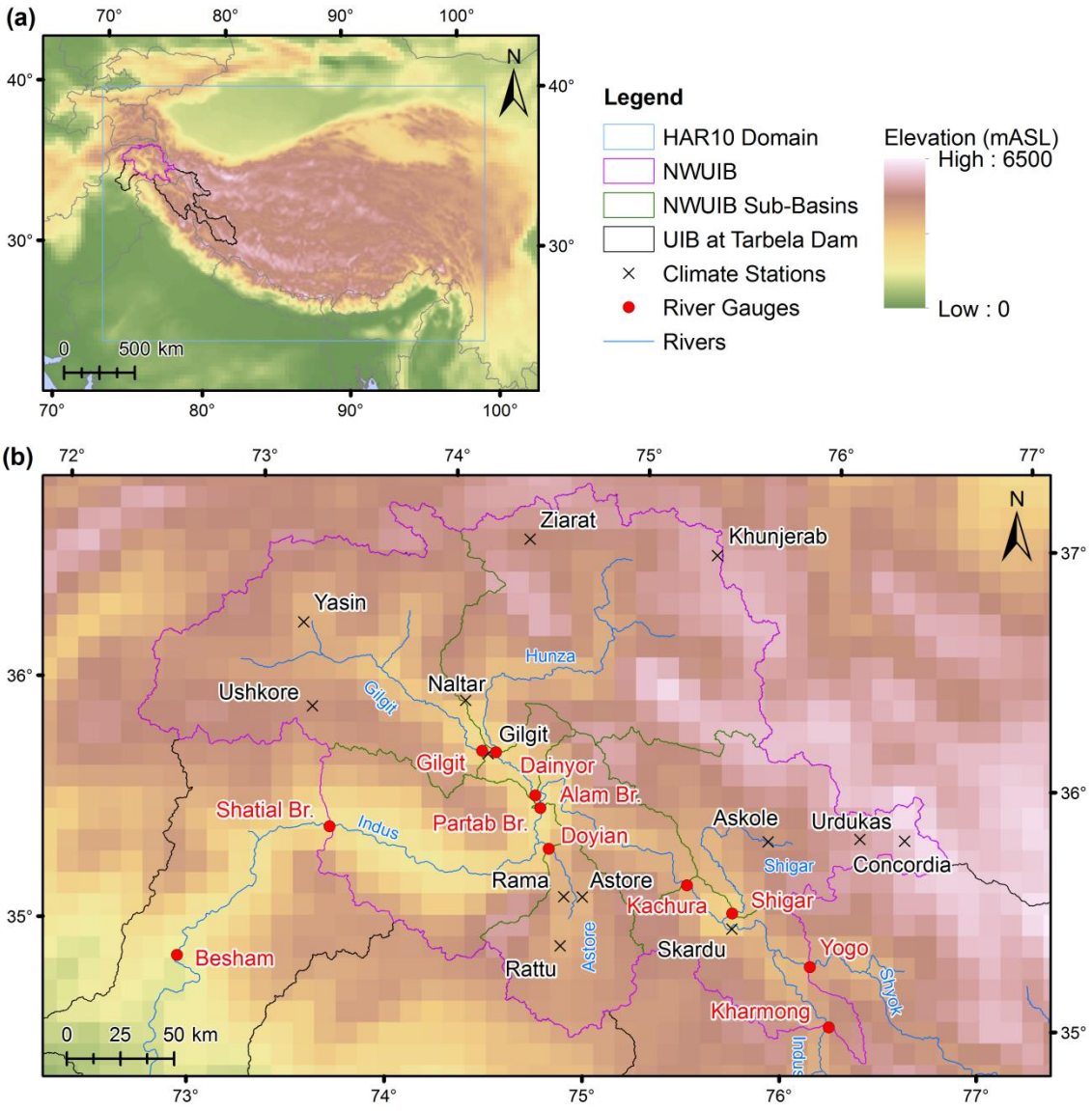


Figure 3.1 – Regional overview and study area maps. The inner HAR domain (HAR10), orography and the UIB study area are given in (a), while (b) indicates the locations of climate stations and river gauges used in this analysis (details in Table 3.2 and Table 3.3). The UIB boundary was delineated by Khan et al. (2014).

Much of the precipitation in the basin occurs in winter and spring, often originating from westerly disturbances (Filippi et al. 2014; Archer and Fowler 2004). These synoptic scale low pressure systems are guided towards the UIB from the Atlantic and Mediterranean by the subtropical westerly jet. There are some relatively infrequent monsoon-related storms in summer, but westerly tracking depressions can also bring precipitation at this time of year, even though the subtropical westerly jet shifts to the north of the Karakoram (Hewitt 2014). Precipitation is strongly orographically forced, leading to a general increase with elevation to around 5500 mASL but then most likely a decrease at higher elevations, due to exhaustion of

moisture availability (Hewitt 2014). The interplay of precipitation sources and topographic barrier effects incurs substantial horizontal gradients in precipitation across the UIB (Young and Hewitt 1990; Reggiani and Rientjes 2015).

Archer (2004) demonstrated that there is high spatial and vertical correlation of observed near-surface air temperatures in the UIB on monthly, seasonal and annual time scales. Temperature measured at a relatively small number of valley-based stations thus explains a substantial proportion of observed variability in melt season runoff in heavily glaciated, energy-limited catchments (Archer 2003; Fowler and Archer 2005). In contrast to many parts of the world, observed summer temperatures in the UIB show a cooling trend (Fowler and Archer 2006; Forsythe et al. 2012b), which is concurrent with comparative glacier mass stability (Hewitt 2005; Brun et al. 2017). This appears to be related to interactions between the summer monsoon and the Karakoram/Western Tibetan Vortex system (Forsythe et al. 2017; Li et al. 2018). Studies of the full surface energy balance in the UIB are limited due to poor data availability, although the modulating influence of cloud cover has been examined by Forsythe et al. (2015).

Here we focus particularly on the north-west part of the UIB (hereafter NWUIB) shown in Figure 3.1. This 52,473 km^2 domain is high-yielding, contributing ~50% of UIB runoff to Besham, upstream of Tarbela Reservoir, despite comprising only ~30% of the catchment area. The NWUIB contains a mixture of the pluvial, nival and glacial hydrological regimes described above.

3.3 Data and Methods

We build our evaluation of the HAR climatology and biases primarily around local hydroclimatic measurements. For reference, we also incorporate some limited comparisons with gridded climate data products commonly applied in this region, namely TRMM and APHRODITE (Yatagai et al. 2012). However, these precipitation products exhibit important limitations in the UIB. Specifically, TRMM does not adequately detect snowfall, which comprises a substantial proportion of UIB precipitation (Forsythe et al. 2012a; Reggiani and Rientjes 2015; Maussion et al. 2014). APHRODITE is based on interpolation of the small number of in-situ observations, which are biased towards drier, valley locations. This inevitably leads to underestimation of precipitation at higher elevations and at the catchment scale (Immerzeel et al. 2015). Therefore, here we focus more on alternative datasets and approaches based on MODIS products, in order to supplement local observations and provide

broader spatial coverages, as detailed in Table 3.1 and Sections 3.3.2-3.3.3. In addition, we do integrate some comparisons with global reanalysis products, drawing on ERA-Interim (Dee et al. 2011), JRA55 (Kobayashi et al. 2015) and NASA MERRA2 (Gelaro et al. 2017). However, our primary aim is to evaluate the high resolution HAR with respect to the best available observations, rather than to compare it with datasets at generally coarser resolutions.

Variable	Comparison	Datasets Used
Precipitation	Point-scale climatology	In-situ observations, ERA-Interim, JRA55, MERRA2
	Catchment-scale water balance closure	Gauged river flows, ERA-Interim, JRA55, MERRA2, APHRODITE, TRMM
	Vertical gradients	Published precipitation-elevation relationships
	Horizontal gradients	Snow water equivalent (SWE) reconstruction based on MODIS remote sensing (MOD10A1 SCA)
Temperature	Point-scale climatology	In-situ observations
	Spatial patterns	In-situ observations (with regression for lapse rates and freezing isotherms)
Humidity	Point-scale climatology	In-situ observations
	Spatial patterns	In-situ observations (vertical profiles)
Radiation	Point-scale climatology	In-situ observations
Cloud	NWUIB spatial climatology	MODIS remote sensing (MOD06L2 cloud cover fraction, CCF)
	Correlation with temperature variability	MOD06L2 CCF and in-situ temperature observations
Albedo	NWUIB spatial climatology	MODIS remote sensing (MCD43A3 albedo)
	Correlation with temperature variability	MCD43A3 albedo and in-situ temperature observations

Table 3.1 – Summary of climate variables investigated, comparisons undertaken and datasets used.

3.3.1 High Asia Refined Analysis (HAR)

The initial and boundary conditions for the WRF simulations comprising the HAR were provided by the NCEP FNL (Final) Operational Global Analysis dataset (ds083.2), which uses the same model as the NCEP Global Forecast System (GFS) (NOAA/NCEP 2000). The

FNL dataset integrates surface and upper air observation networks, as well as remote sensing products. The HAR used a daily re-initialisation strategy to prevent drift from observed synoptic conditions over its total simulation period (October 2000 to October 2014). Each day was simulated independently of all other days. For any given day, 36 hours were simulated, beginning at 12:00 UTC on the previous day. The first 12 hours were then discarded as a spin-up to decrease the influence of the initial conditions, which were interpolated from the FNL dataset to the model grid. The outputs of all of the individual daily simulations were ultimately concatenated to produce the overall time series.

This approach to initialisation and spin-up was found to perform well relative to other computationally feasible strategies during extensive testing (Maussion et al. 2011, 2014) and in other Himalayan modelling studies (e.g. Norris et al. 2015; Cannon et al. 2017b). However, it does lead to a degree of discontinuity in the time series for key land surface states, such as snow water equivalent (SWE). This has implications for the land surface water balance, as discussed in Sections 3.3.2 and Appendix A, Section A1.3.

Two-way nesting was applied to run WRF with a 10 km resolution child domain (HAR10) within a broader 30 km resolution parent domain. The 30 km domain was then run separately, without the child domain, to remove any inconsistencies, thereby producing HAR30. Differences between HAR10 and HAR30 therefore provide insights into the effects of resolution and simulation strategy. Further details are given in Appendix A (Section A1.1) and Maussion et al. (2014).

3.3.2 In-situ Observations and Data Processing

Climate Observations

This study uses local climate observations from 13 stations covering a range of elevations, as detailed in Table 3.2 (see also Figure 3.1b). These datasets have been subjected to quality control procedures and used in a number of previous studies (e.g. Archer and Fowler 2004; Fowler and Archer 2006; Collier et al. 2013; Soncini et al. 2014). We performed additional checks for inconsistencies, spurious values and outliers, particularly for the most recent parts of the data. For a given station, these checks included comparisons of recently obtained data with climatologies derived from earlier parts of the time series, as well as inspections of inter-station and inter-variable consistency.

ID	Station	Latitude	Longitude	Elevation (mASL)	Variables	Period	Source
1	Astore	35.366	74.900	2394	P,T	1961-2010	PMD
2	Gilgit	35.916	74.333	1460	P,T	1961-2010	PMD
3	Skardu	35.300	75.683	2210	P,T	1961-2010	PMD
4	Rama	35.358	74.805	3140	P,T	1995-2008	WAPDA
5	Rattu	35.152	74.816	2920	P,T	1995-2008	WAPDA
6	Yasin	36.369	73.300	3353	P,T,RH	1995-2008	WAPDA
7	Ushkore	36.025	73.400	3353	P,T,RH	1995-2008	WAPDA
8	Naltar	36.127	74.184	2810	P,T,RH	1995-2012	WAPDA
9	Ziarat	74.418	36.829	3688	P,T,RH	1995-2012	WAPDA
10	Khunjerab	36.850	75.400	4733	P,T,RH	1995-2012	WAPDA
11	Askole	35.681	75.815	3015	T,RH,SW	2005-2014	EvK2CNR
12	Urdukas	35.728	76.286	3926	T,RH,SW	2004-2014	EvK2CNR
13	Concordia	35.744	76.514	4700	T,RH,SW,LW	2012-2014	EvK2CNR

Table 3.2 – Climate stations used in this study. Variable name abbreviations are: P – precipitation, T – temperature, RH – relative humidity, SW – incoming shortwave radiation, LW – incoming longwave radiation. Data source abbreviations are: PMD – Pakistan Meteorological Department, WAPDA – Water Power and Development Authority, EvK2CNR – Everest-K2-Consiglio Nazionale delle Ricerche (research group).

To compare the HAR with climate observations, we extracted hourly time series from the closest HAR grid point to each station location, similarly to Maussion et al. (2014). In the analyses, we include only days for which observations are available and months that are largely complete (less than 3 days missing data). For the evaluations of monthly temperature biases (Section 3.4.2, Figure 3.6a) and incoming longwave radiation (Section 3.4.4, Figure 3.10), we apply corrections for elevation differences between stations and HAR grid points. The corrections are based on local, monthly climatological gradients (i.e. lapse rates), which were calculated from the 9 HAR grid cells surrounding the station location using linear regression. The elevation correction procedure reduced mean annual temperature and longwave radiation biases by 4.6°C and 12 W/m^2, respectively. Details of the regressions and

sensitivity of the elevation correction to different timescales are discussed in Section A1.2 of Appendix A. We did not make similar elevation corrections in the other evaluations in Section 3.4, due to the nature of the comparisons we employ, as well as the weaker, less consistent and more complicated dependence of the other variables on elevation.

Limitations here include the difference in scale between weather stations and the HAR grid resolution, as well as uncertainties inherited from numerical approximations at grid cells. Where applied, the elevation correction also omits the important influence of more local aspect and slope variation (Daly et al. 2008). This is difficult to account for given the low station density and the limitations of the HAR orography. With these issues in mind, we compare the datasets at multiple locations in relative as well as absolute terms, generally at monthly or longer time scales to reduce noise and identify robust patterns.

Hydrological Observations

We consider the plausibility of HAR precipitation at the catchment scale based on river flow records for 11 gauged sub-basins of the UIB (Table 3.3 and Figure 3.1b). These data have been checked and applied in previous studies (Archer 2003; Sharif et al. 2013). We calculated HAR catchment means accounting for partial grid cell coverage, focusing on mean annual aggregations to diminish any confounding influences from inter-annual storage changes. We use overlapping periods of record as far as possible. In addition to comparing HAR precipitation with observed runoff directly, we also derived 'effective precipitation', defined here as precipitation minus evapotranspiration. This was calculated based on mean annual aggregations, with both terms taken from the HAR. Effective precipitation should be approximately equal to runoff in the mean annual case if storage changes are close to zero, an assumption which is consistent with the neutral glacier mass balances in much of the UIB (Hewitt 2005; Zhou et al. 2017; Bolch et al. 2017; Brun et al. 2017).

The effective precipitation approach was taken primarily because the runoff simulated by the Noah LSM in the HAR was found to be strongly affected by the limited accuracy of daily SWE re-initialisation from the coarse FNL driving dataset. Limitations of our approach include notable uncertainty in HAR evapotranspiration, which means that this analysis should be considered indicative. These issues are discussed in Section 3.4.1 and Appendix A, Section A1.3.

ID	Station	River	Latitude	Longitude	Area (km^2)	Period
1	Kharmong	Indus	34.933	76.228	70,030	1982-1997
2	Yogo	Shyok	35.176	76.100	33,157	1973-2008
3	Shigar	Shigar	35.366	75.676	7,040	1985-1987
4	Kachura	Indus	35.463	75.428	114,093	1970-2008
5	Dainyor	Hunza	35.926	74.371	13,732	1966-2012
6	Gilgit	Gilgit	35.925	74.300	12,671	1961-2008
7	Alam Bridge	Gilgit	35.766	74.600	27,272	1966-2008
8	Partab Bridge	Indus	35.716	74.633	144,407	1962-2008
9	Doyian	Astore	35.550	74.700	3,899	1974-2013
10	Shatial Bridge	Indus	35.533	73.566	155,689	1983-1997
11	Besham	Indus	34.916	72.875	164,867	1969-2013

Table 3.3 – Details of WAPDA river flow gauges used in this study.

3.3.3 MODIS Products and Processing

Snow Water Equivalent (SWE) Reconstruction

Using the Collection 6 MOD10A1 500 m daily snow cover product (Hall and Riggs 2016), with cloud gaps infilled using the method of Gafurov and Bárdossy (2009), we employed a well-established procedure for reconstructing spatially distributed peak SWE (e.g. Raleigh and Lundquist 2012). In this approach, the timing of snow disappearance is identified from remote sensing, and snowmelt prior to this date is calculated using a model. In line with other studies, we use a simple temperature-index (degree day) algorithm (e.g. Rice et al. 2011; Raleigh and Lundquist 2012). This is driven by observed daily temperature time series, adjusted for elevation with monthly climatological lapse rates estimated from linear regression using station data and elevations. The purpose of reconstructing peak SWE in this way is to gain a proxy for relative spatial variation in UIB mass input to evaluate the HAR's spatial and vertical precipitation gradients.

This method is predicated on the previous finding that air temperature provides a good index of energy available for ablation in this region (Archer 2004). However, its applicability is limited to areas where melting is mass-constrained, i.e. locations that are not glacierised or perennial snowpacks. Elsewhere, the primary assumption is that patterns of relative spatial variation in SWE are roughly consistent with those of precipitation at fairly broad scales. There are of course methodological limitations, including the simplification of temperature variability and its relationship with ablation, as well as the omission of snow redistribution and sub-grid patterns. Nevertheless, our primary goal is only to indicate likely relative variation in mean annual mass input at much broader scales than the MODIS pixels.

Cloud Cover

The Collection 5.1 MOD06L2 daily 5 km cloud cover fraction (CCF) product (Platnick et al. 2003) is used to understand some of the patterns identified in the HAR. The applicability of this dataset in the UIB was demonstrated by Forsythe et al. (2015). We use MOD06L2 spatial means of daily average CCF for the NWUIB, as well as time series of the means of the 9 pixels in MOD06L2 surrounding each climate station location. The equivalent HAR cloud cover variable was calculated by Maussion et al. (2014) as the maximum CCF in a 50 km horizontal view field for each cell, based on classifying all model layers using a threshold condensate mixing ratio, following Mölg and Kaser (2011). While there are differences in the provenance of MODIS and HAR CCF, we take these datasets as indicative of cloud cover patterns at generally coarse aggregations.

Albedo

We also use spatial means and station location time series from the Collection 5 MCD43A3 albedo product (shortwave band). The accuracy of this product has been established previously (e.g. Wang et al. 2012), although the challenges in albedo retrieval do increase in complex terrain (Wen et al. 2018). As such, we use the snow albedo from the MOD10A1 500 m daily snow cover product (Hall and Riggs 2016) for comparison, infilling no-snow pixels using MCD43A3 white sky albedo to provide an upper estimate of albedo.

3.4 Results

3.4.1 Precipitation

Point-Scale Comparison

For a first indication of precipitation performance, we compare HAR mean annual values with observations (Figure 3.2a). This shows that HAR10 is generally consistent with most observations, while HAR30 tends to overestimate relative to station records. The exceptions to good HAR10 agreement with measurements are three relatively high elevation stations – Khunjerab, Ziarat and Ushkore – for which the HAR simulates high annual totals. This could be due to HAR limitations. For example, reductions in precipitation downwind of the major topographic barriers to the south and east of Khunjerab and Ziarat could be underestimated (see Figure 3.1). However, the Khunjerab station is also located on an exposed, high mountain pass, such that the influences of strong local wind patterns unresolved by the HAR may be especially important. Ushkore and Ziarat both lie in relatively narrow valleys that are not captured in the HAR orography, such that local, valley-scale vertical precipitation gradients are not simulated in these cases. Figure 3.1b shows that most of the other stations recording precipitation sit in major valleys that are at least partly represented. These points highlight the challenge of scale differences for precipitation evaluation at station locations, while measurement limitations must also be acknowledged. As such, we augment our analysis with appraisal of precipitation gradients and sub-basin water balances below.

To examine the intra-annual distribution of precipitation, we show the fractions of annual precipitation occurring in each month for the station ensemble in Figure 3.2b. From this we can see that the HAR essentially agrees with the observed annual cycle of precipitation, with a maximum in winter/spring and a minimum in summer. However, the HAR generally overestimates the proportion of precipitation falling in winter and early spring, while underestimating the summer fraction. Figure 3.2b also shows that the HAR10 and HAR30 cycles are very similar, indicating little modification of precipitation seasonality by grid resolution. Notably, the observations exhibit more spatial variation in the shape of the annual precipitation cycle than exists in the HAR.

Interestingly, Reggiani et al. (2017) show that some of the coarser resolution global reanalyses, including ERA-Interim and JRA55, have a less stark annual cycle, which is more consistent with observations (see also Appendix A, Section A2.1). This raises the possibility

that HAR seasonality could be improved with alternative boundary conditions, although it is also possible that the limitations stem from aspects of the HAR setup, for example parameterised convection. Additional point-scale comparisons are given in Appendix A (Section A2), including time series, which support Maussion et al.'s (2014) finding that the HAR reproduces precipitation inter-annual variability.

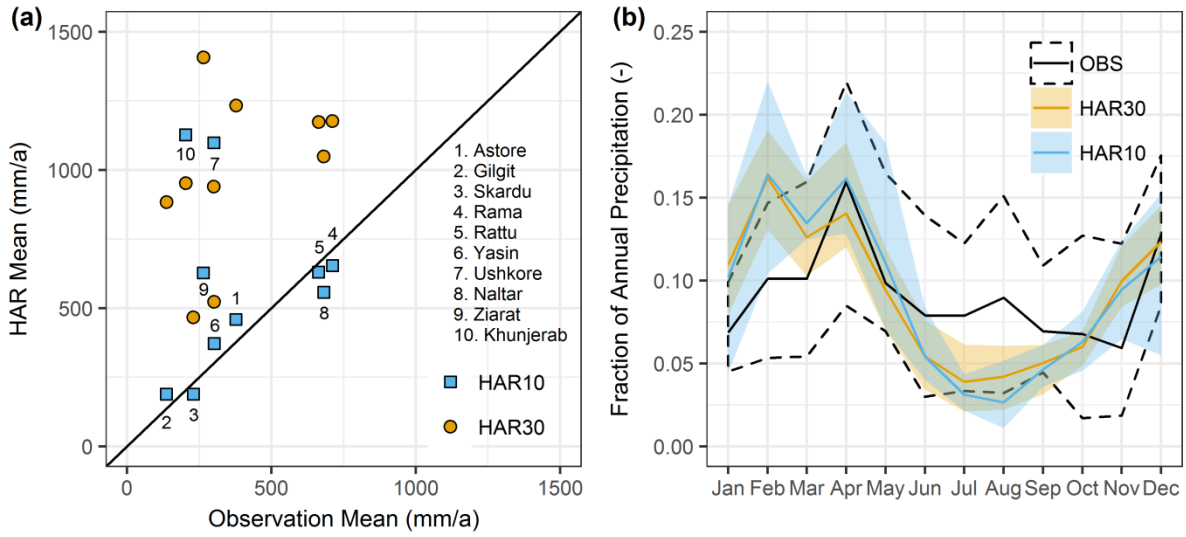


Figure 3.2 – Comparison of HAR precipitation with station observations (OBS). Annual means are given in (a), with numbers identifying stations given next to HAR10 points, while (b) shows mean monthly precipitation normalised to mean annual totals for the station ensemble. The solid line in (b) shows the mean for all stations, while shading shows the range.

Catchment-Scale Assessment

Moving to the catchment scale, we first compare spatially averaged mean annual precipitation from the HAR with observed runoff for gauged sub-basins of the UIB (Figure 3.3a). The results indicate that the HAR reproduces the intuitively positive relationship between mean annual precipitation and observed runoff. It is also apparent that the simulated precipitation exceeds observed runoff for all catchments, although the nival Astore and Gilgit sub-basins do approach parity. Based on HAR10, we find the ratio of annual runoff to precipitation to be approximately 0.65. In addition, Table 3.4 shows that HAR annual precipitation for the UIB to Besham, upstream of Tarbela Reservoir, is similar to the ERA-Interim, JRA55 and MERRA2 global reanalyses, with all estimates falling in the range of 721 to 793 mm/a for the period considered. In contrast, precipitation from the APHRODITE and TRMM products is substantially lower than observed runoff. This highlights the limits to their applicability in the UIB, as discussed in Section 3.3.

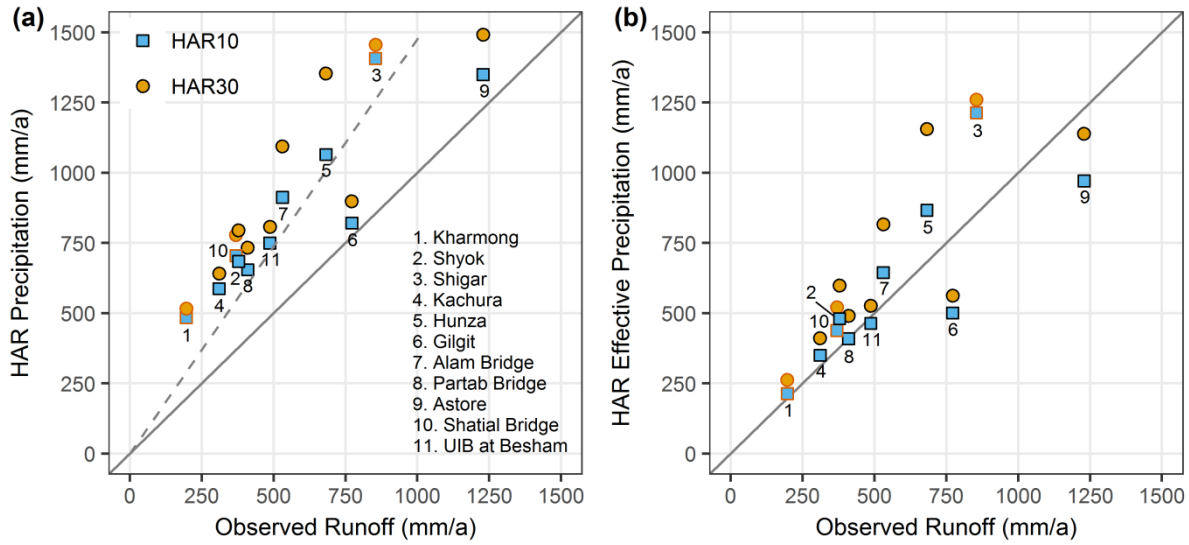


Figure 3.3 – Comparison of HAR precipitation and water balance with observed runoff. HAR mean annual precipitation is compared with observed runoff in (a), while HAR effective precipitation is compared with observed runoff in (b). The dashed line in (a) is from regression, allowing estimation of a runoff coefficient. Effective precipitation in (b) is calculated by subtracting HAR-simulated evapotranspiration from precipitation. Red outlines denote stations without data overlapping the HAR period, such that only the available record period was used. Numbers identifying gauges are given next to HAR10 points.

We examine the HAR precipitation further by accounting for evapotranspiration, as simulated by the HAR. The resulting quantity is termed effective precipitation here, which should be approximately equal to runoff at the mean annual scale considered, if storage changes are close to zero. Figure 3.3b shows that HAR mean annual effective precipitation and observed runoff do indeed generally converge, particularly for low- to moderate-yielding sub-basins. For higher-yielding sub-basins, larger errors are evident, although there does not appear to be systematic over- or under-estimation if all sub-basins are considered together, at least for HAR10.

This suggests that the HAR may provide a reasonable approximation of catchment-scale precipitation on the whole, but the strength of this finding of course depends on the accuracy of HAR-simulated evapotranspiration and assumptions about storage changes. With glaciers being the major source of storage in the UIB, the storage change assumption is consistent with the approximately neutral glacier mass balances observed in the heavily glaciated parts of basin (Hewitt 2005; Zhou et al. 2017; Bolch et al. 2017; Brun et al. 2017). HAR mean annual evapotranspiration is approximately 275 mm for the UIB to its outlet at Besham, ranging from around 200 to 400 mm between sub-basins. These evapotranspiration values are within the

range of estimates from reanalyses (see Table 3.4), remote sensing and hydrological modelling used in other recent studies (Immerzeel et al. 2015; Reggiani et al. 2017; Reggiani and Rientjes 2015).

Data Source	P (mm/a)	ET (mm/a)	P - ET (mm/a)	Qobs (mm/a)	P - Qobs (mm/a)	P - ET - Qobs (mm/a)
APHRODITE	215	n/a	n/a		-246	n/a
ERAI	735	336	399		274	-62
HAR10	721	276	445		260	-16
HAR30	778	275	503		317	42
JRA55	793	127	667		332	206
MERRA2	724	184	540	461	263	79
TRMM 3B42	288	n/a	n/a		-173	n/a
TRMM 3B43	318	n/a	n/a		-143	n/a
Reggiani and Rientjes (2015)	675 ± 100	200 ± 100	n/a		n/a	n/a

Table 3.4 – Comparison of HAR mean annual water balance for the UIB at Besham with other gridded data products (based on overlapping record period of October 2000 to September 2007). Abbreviations are for precipitation (P), evapotranspiration (ET), and observed runoff (Qobs), with "-" representing the minus sign. APHRODITE and TRMM data products do not include evapotranspiration.

While consistency with other data sources is useful, clearly it does not provide complete validation of the HAR's absolute accuracy for evapotranspiration. Substantial uncertainties due to various errors and biases are of course associated with both HAR and alternative estimates. Although this means that the analysis should be considered indicative, it does at least suggest that the HAR shows notable potential as a source of catchment-scale precipitation for the data-sparse UIB, according to the best sources of information we have on other water balance terms at present (see also Appendix A, Section A1.3). Further analyses through process-based hydrological modelling could explore this in more detail. Moreover, we can conclude with some confidence that HAR precipitation outperforms APHRODITE and TRMM at the catchment scale. How much information it adds relative to global

reanalyses depends in part on its vertical gradients and patterns of spatial variation, as explored in the next section.

Vertical Gradients and Spatial Variation

We now examine whether the HAR simulates the vertical precipitation gradients required to accurately reproduce cryospheric accumulation and ablation patterns in the UIB, and therefore runoff amounts and timing (Hewitt 2014). The HAR hypsometry is provided in Figure 3.4a, demonstrating the substantial benefit of a 10 km resolution, while Figure 3.4b shows the precipitation-elevation relationship for the NWUIB (based on HAR10 only). This relationship takes a log-linear form. The points in Figure 3.4b, each representing a HAR10 grid cell, are coloured according to the standardised residuals calculated after applying regression to the HAR10 precipitation-elevation relationship. Two observation-based profiles of precipitation are also shown on Figure 3.4b (Hewitt 2014; Winiger et al. 2005), while additional profiles from model-based precipitation reconstructions are given in the Appendix A (Section A2.5).

Importantly, Figure 3.4b suggests that the central tendency of the HAR10 vertical profile agrees quite closely with Winiger et al. (2005)'s profile, the formulation of which has been the basis for several studies of UIB hydroclimatology (e.g. Bocchiola et al. 2011; Soncini et al. 2014; Reggiani et al. 2016). Some contrasts with the observation-based profiles are of course expected, based on the differing and restricted spatial extents of the measurement networks used, for example. In addition, field observations suggest that precipitation may begin to decrease at very high elevations (above around 5000-5500 mASL) due to exhaustion of moisture availability, as noted in Section 3.2 (e.g. Hewitt 2014). However, Figure 3.4b indicates that HAR10 does not show a decrease in precipitation at its highest elevations. This may be related to topographic smoothing in HAR10, as its elevation peaks at around the same point at which the precipitation inversion is expected to start. This could mean that a decrease is simply not visible or is potentially inhibited by the effects of topographic smoothing on orographic precipitation dynamics and thermodynamics (Cannon et al. 2017b).

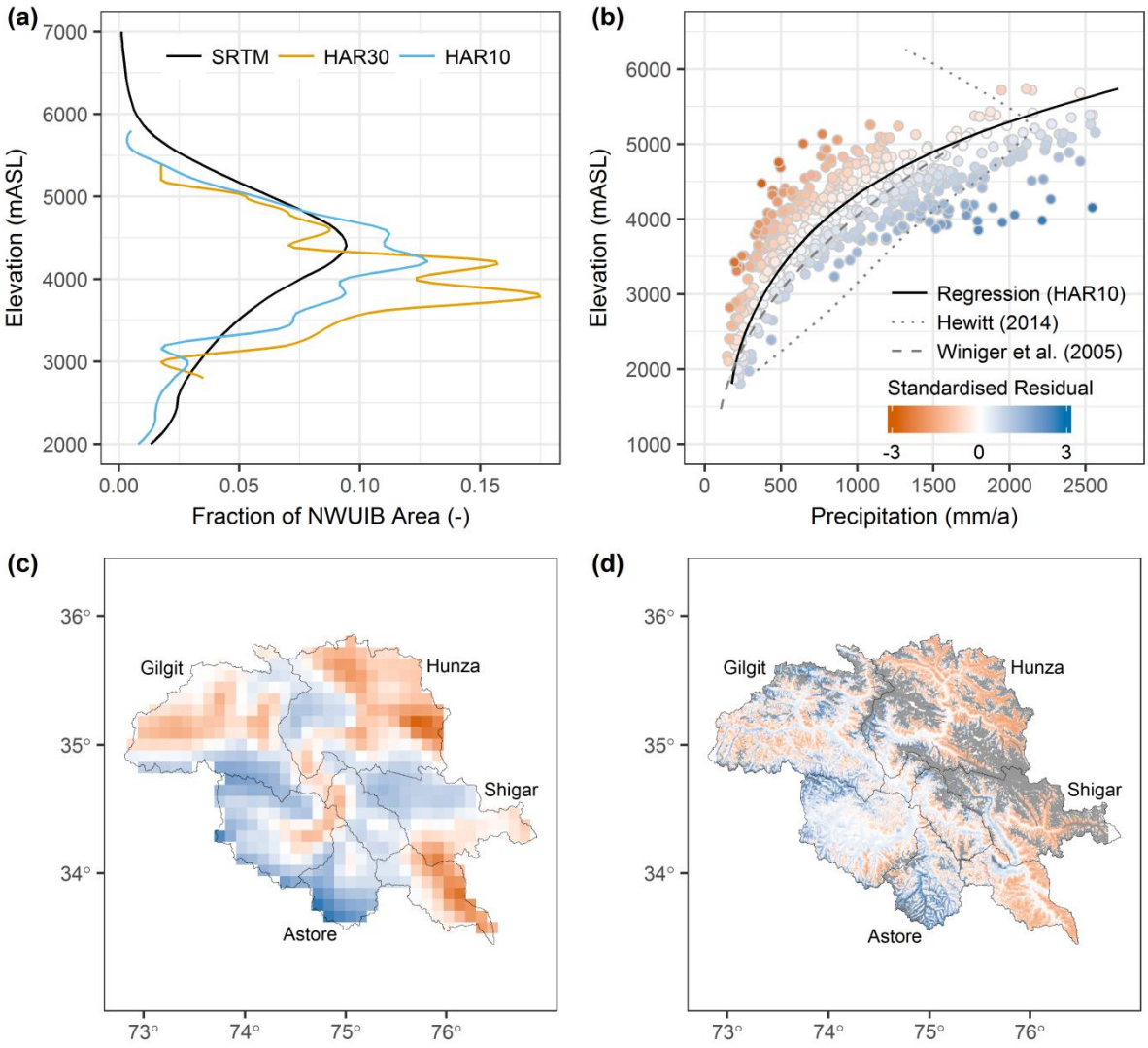


Figure 3.4 – Spatial and vertical precipitation gradients, compared with observation-based profiles and relative patterns of SWE reconstructed from remote sensing. HAR hypsometry is compared with the Shuttle Radar Topography Mission (SRTM) reference dataset in (a), while (b) shows a scatterplot of mean annual precipitation and elevation for each HAR10 cell in the NWUIB. The black line follows a log-linear profile determined from regression, and the points are coloured according to standardised residuals from this relationship. Two observation-based vertical precipitation profiles are also plotted (Hewitt 2014; Winiger et al. 2005). The spatial distribution of HAR standardised residuals from the precipitation-elevation relationship in (b) is shown in (c), using the same colour scale. The spatial distribution of standardised residuals after removing the elevation signal from (MODIS) reconstructed SWE is given in (d), using the same colour scale as (b) and (c). Glaciers are shown in (d) in grey.

To understand the notable scatter in Figure 3.4b, we map the cell-wise standardised residuals back to the NWUIB domain in Figure 3.4c to reveal their patterns of spatial correlation. The red and blue areas can be interpreted as showing which parts of the NWUIB have relatively low or high precipitation, respectively, compared with what might be expected from elevation

alone (i.e. black line in Figure 3.4b). This allows us to see spatial variation at the sub-regional scale, but clearly scales finer than the HAR10 grid are unresolved. From Figure 3.4c we can see relatively high precipitation zones along the west and south-west margins of the NWUIB, the areas of the domain first encountered by the prevailing westerly flows. This is consistent with the high yield of the nival Astore sub-basin (Figure 3.3). A second zone of relatively high precipitation trending north-west to south-east (from the Hunza into the Shigar sub-basin) can be seen to follow the terrain (see Figure 3.1b). To the north-west of this ridge in the Hunza basin, precipitation is comparatively low for its elevation, suggesting that the known rain shadow effect here is captured by the HAR to some degree at least (Winiger et al. 2005). Other prominent zones of relatively low precipitation and vertical gradients are found in the central Gilgit catchment, and the south-east corner of the NWUIB adjoining the Shigar sub-basin (see also Figure 3.1b).

For qualitative evaluation of these patterns, we compare the HAR with reconstructed SWE (Section 3.3.3). Similar to Figure 3.4c, we show a spatial distribution of standardised residuals in Figure 3.4d, but this time based on the regression of reconstructed SWE and elevation, rather than the regression of HAR precipitation and elevation that underpins Figure 3.4c. Glaciers are also shown, as SWE reconstruction is not undertaken in these areas. Comparing Figure 3.4c and Figure 3.4d shows several similarities, such as the relatively high precipitation amounts in the Astore sub-basin and around the southern margin of the Gilgit sub-basin. It is clear that the major band of glacierised area, running north-west to south-east from the Hunza into the Shigar sub-basin in Figure 3.4d, corresponds with a relatively high precipitation zone in the HAR. There is also agreement on the zones of comparatively low precipitation relative to their elevations, including the central part of the Gilgit sub-basin, but particularly the south-east and north-east corners of the NWUIB.

There are of course limitations in this qualitative comparison. In particular, it presumes high correlation of precipitation, snowfall and peak SWE, but this is expected to be generally the case at the scale of multiple HAR cells discussed here (Appendix A, Section A2.5). Thus, despite the uncertainties, the key similarities between Figure 3.4c and Figure 3.4d suggest that the HAR patterns of relative spatial variation in precipitation are physically consistent with what we can infer from remote sensing. In conjunction with the generally reasonable agreement with both climate and hydrological observations, this analysis suggests that the HAR10 reproduces some critical features of UIB precipitation amounts and spatiotemporal distribution. The HAR10 may thus approach a resolution high enough to capture the most

important effects of orographic precipitation in this area, particularly given the dominance of westerly winter/spring events (Norris et al. 2016; Daly et al. 2017).

3.4.2 Temperature

The annual cycle of monthly mean near-surface (2 m) air temperature for the station ensemble is shown in Figure 3.5a. Prior to calculating the ensemble summary statistics, temperatures for each station were expressed as differences relative to their annual means, in order to remove the effects of elevation differences between stations. This provides a way to examine overall differences in the shape of the annual cycles. We can therefore see from Figure 3.5a that HAR10 and HAR30 have very similar annual cycles and both agree with observations in general. However, the amplitude of the cycle is overestimated by the HAR, mainly through a higher summer peak relative to the observed cycle. The rate of temperature increase in spring and early summer is also initially shallower than for observations, before becoming comparatively steeper.

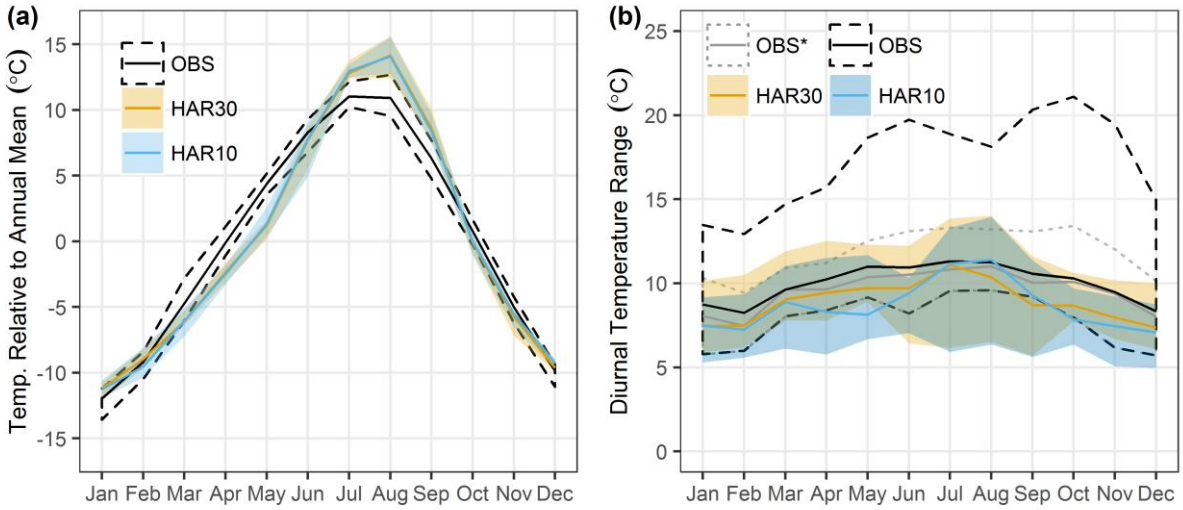


Figure 3.5 – Comparison of HAR annual temperature cycles with observations. Mean monthly temperatures are shown in (a) and diurnal temperature range (DTR) in (b). Solid lines show medians, whereas shaded areas and dashed lines show ranges across all stations. The mean monthly temperatures in (a) were first normalised for each station individually, by subtracting the respective annual mean. All stations were then used to compute the ensemble summary statistics. The OBS* profile in (b) omits two stations with particularly high DTR, Gilgit and Skardu, from the ensemble.

Figure 3.5b displays the annual cycle of monthly mean diurnal temperature range (DTR), an important variable that contains signals from cloud influences on radiation and other factors

affecting near-surface heating and fluxes. Figure 3.5b demonstrates that observed DTR is typically higher than in the HAR. However, it does also show that the HAR captures the comparatively modest annual DTR cycle for most stations, with DTR highest in the summer months. On average, the HAR tends to be most biased in spring, early summer and autumn, underestimating DTR at these times, especially in HAR10. The largest underestimation of DTR occurs for lower elevation stations, especially Gilgit and Skardu, which could reflect a cold pool effect not present in the HAR.

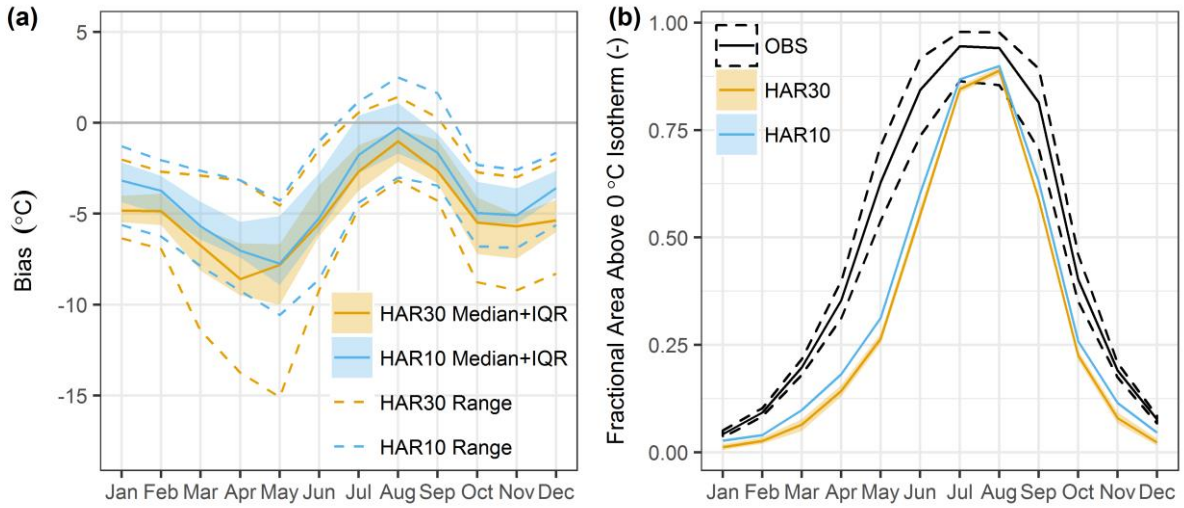


Figure 3.6 – HAR temperature bias and its hydrological implications. Differences between elevation-adjusted HAR monthly mean temperatures and station observations are shown in (a). The median, interquartile range and the range of differences across all stations are given. The fraction of NWUIB area above the freezing isotherm according to linear regression is illustrated in (b), for both observations and the HAR (ranges show 95% confidence intervals). HAR temperatures were adjusted for elevation differences compared with stations, using local lapse rates in the HAR, identified through regression.

Using the HAR elevation-corrected temperatures to examine the bias in the monthly means (Section 3.3.2) reveals that the HAR is generally colder than observations (Figure 3.6a). The bias is largest in spring and smallest in summer, with HAR30 generally being slightly colder than HAR10. In Appendix A (Section A1.2), we show that this annual cycle of bias is robust to the alternative methods of elevation correction tested. Importantly, the peak cold bias in spring coincides with the largest underestimation in the temperature cycle shown in Figure 3.5a, suggesting that the HAR does not warm quickly enough at this time of year. The smallest temperature bias in summer occurs when the HAR shows a more rapid and pronounced increase in the annual temperature cycle than observations (Figure 3.5a). Controls

on rates of change in the annual temperature cycle are therefore likely to be critical to the seasonally varying magnitude of the cold bias in the HAR.

To assess the implications of this temperature bias for runoff, we use linear regression of temperature and elevation to examine the proportion of the NWUIB area that lies above the freezing isotherm each month. This provides a broad indication of the area of the catchment where snow and ice melt can take place, termed here 'melt area'. Figure 3.6b shows that the HAR has a substantially compressed annual cycle of melt area compared with what we can infer from observations. As melt area is a key control on runoff sensitivity and variability in the UIB (Forsythe et al. 2012b), understanding the causes and implications of this bias is important for hydrological applications, as well as for minimising biases in future WRF simulations. Additional analysis of temperature variability and lapse rates, both fairly well simulated, is given in Appendix A (Section A3).

3.4.3 Humidity

Figure 3.7a depicts the annual cycle of specific humidity variation as monthly differences relative to the annual mean, similarly to Figure 3.6a. There is generally good agreement between the observed and HAR datasets, although the HAR appears to underestimate the amplitude of the annual cycle. In particular, it shows a higher winter minimum than observations, but a lower summer peak. Expressed as relative humidity, the HAR provides an overestimate in winter, spring and autumn, before converging on observations in summer (Figure 3.7b). The HAR thus displays a comparatively pronounced annual relative humidity cycle. The closer match in summer in relative humidity coincides with the dampened peak in the specific humidity cycle, as well as lower apparent temperature bias at this time of year (see Section 3.4.2). Conversely, overestimation of the specific humidity minimum in winter – and relative humidity in winter, spring and autumn – occurs during periods of colder temperature bias.

To see whether the HAR captures spatial and vertical variation in humidity, we plot seasonal elevation profiles for winter (December, January, February – DJF) and summer (June, July, August – JJA) in Figure 3.8a and Figure 3.8b, respectively. The strong elevation-dependence of specific humidity in winter is well simulated, although there may be some underestimation at lower elevations; more local observations would be required to demonstrate this. In summer, the HAR shows a substantial increase in spatial variation of specific humidity, which is also present in observations. This means the observations are likely to be less spatially

representative, but in general the level of consistency is good. Equivalent vertical profiles for relative humidity are given in Appendix A (Section A4), which shows better consistency with observations is present in summer.

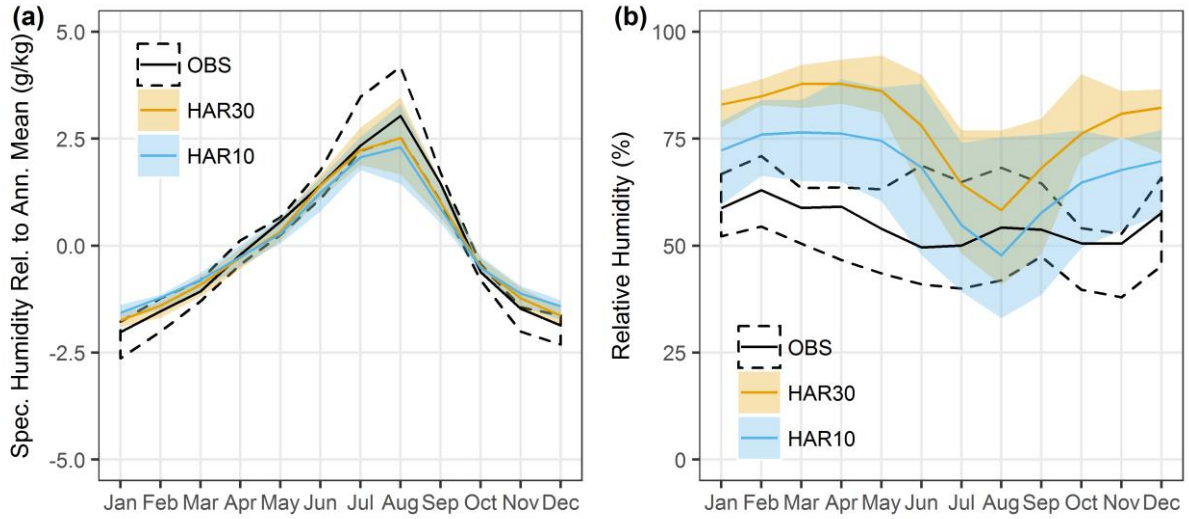


Figure 3.7 – Comparison of HAR humidity with station observations. Monthly specific humidity normalised by subtracting the annual mean is given in (a), while (b) displays monthly relative humidity. The median and ranges across stations are shown.

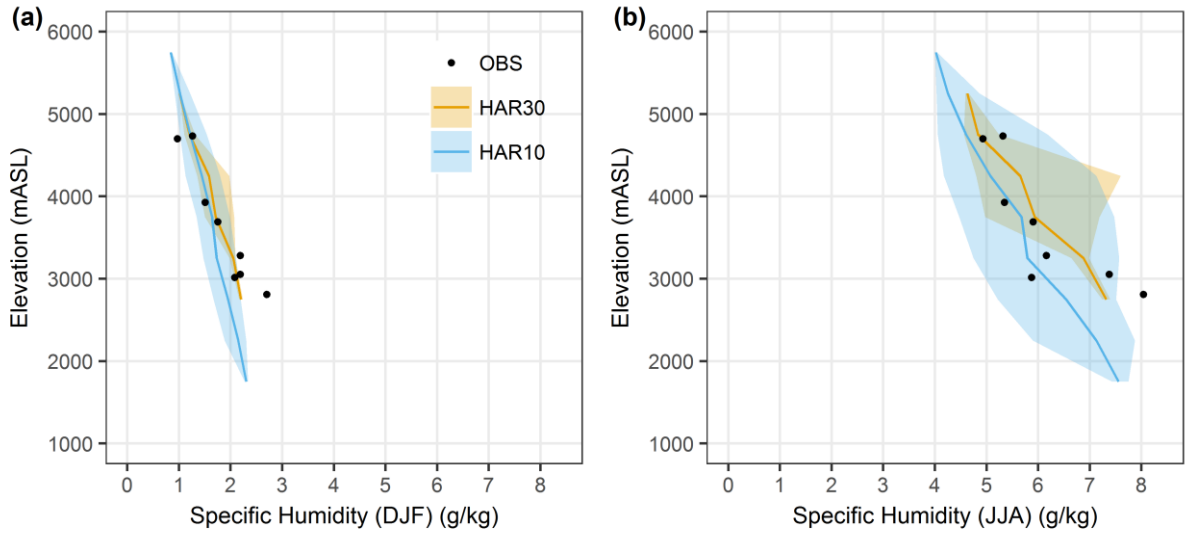


Figure 3.8 – Comparison of observed and HAR vertical profiles of specific humidity for (a) winter (DJF) and (b) summer (JJA). HAR profiles show the elevation band means (solid lines) and ranges (shading).

There are several hydrological implications from these humidity patterns. Firstly, high relative humidity could lead to a low vapour pressure deficit, suppressing evapotranspiration calculated using the FAO Penman-Monteith approach for example (Allen et al. 1998). However, the relative humidity (and temperature) bias generally reduces in summer when evapotranspiration is most significant in the UIB (Reggiani et al. 2017), thereby diminishing the magnitude of this issue, although further simulations would be required to quantify this more precisely. Secondly, as specific humidity is reasonably well represented, it may provide useful direct inputs to snow and glacier surface energy balance models, where latent heat fluxes depend on near-surface gradients in specific humidity or vapour pressure.

3.4.4 Incoming Radiation

In Figure 3.9 we compare HAR incoming shortwave radiation time series (daily and 28-day moving average) with observations from the Askole EvK2CNR AWS. The daily time series demonstrates that the HAR and observed peaks match well throughout the year. This confirms that the HAR accurately simulates incoming shortwave radiation under clear-sky conditions. In contrast, the observed moving average series are clearly lower than the HAR, particularly in summer. This is also the case for the two other EvK2CNR AWSs, Urdukas and Concordia, as well as for other years where data are available (Appendix A, Section A5). Given the good level of agreement for incoming shortwave peaks, i.e. clear-sky conditions, the differences appear to be related to the underestimation of cloudiness or cloud reflection effects. This is particularly noticeable in the HAR daily time series in Figure 3.9, where the magnitude and frequency of cloud effects are visibly underestimated in summer.

For incoming longwave radiation, we compare the HAR with observations from the Concordia AWS in Figure 3.10. The daily time series in this figure shows that in general the HAR captures quite accurately the observed peaks in incoming longwave radiation for much of the year. However, the HAR displays an overall underestimation of longwave radiation, particularly in summer. The magnitude of daily variability relative to observations is also too large in spring, summer and autumn. During clear-sky conditions the HAR, especially HAR10 in fact, may therefore underestimate incoming longwave radiation, while during cloudy conditions it may induce more longwave enhancement than apparent in observations. However, longer time series of in-situ observations from multiple stations would be required to confirm the generality of this finding.

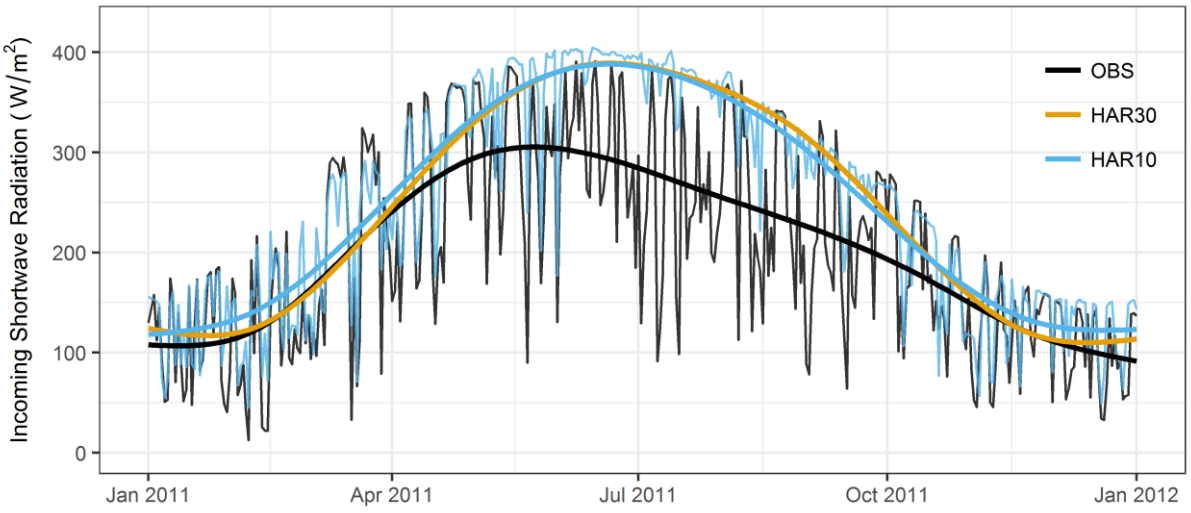


Figure 3.9 – Time series of incoming shortwave radiation for Askole AWS, showing observations and corresponding HAR cell. Thin lines show daily time series, and bold lines show 28-day moving averages. For clarity, only HAR10 daily time series are plotted.

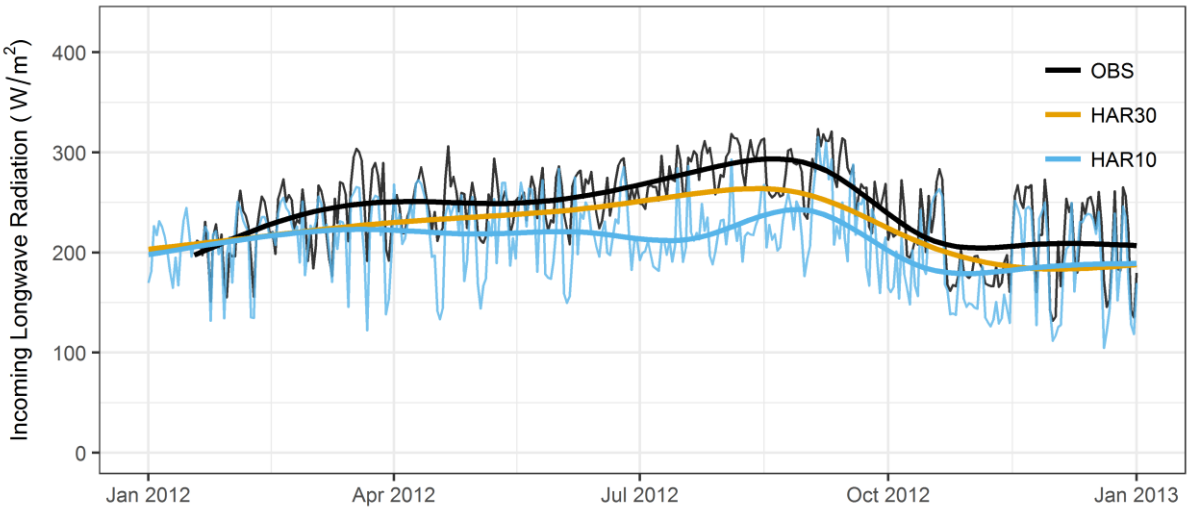


Figure 3.10 – As Figure 3.9 but showing incoming longwave radiation at the Concordia AWS.

Therefore, in late spring, summer and early autumn, the underestimation of cloudiness suggested by the incoming shortwave radiation comparison may correspond to a low bias in incoming longwave radiation. The bias towards too high a frequency of clear-sky conditions peaks in summer, which would be expected to lead to a positive bias in total incoming radiation, due to the higher magnitude of the shortwave flux at this time of year.

3.4.5 Cloud and Albedo Influences

Climatology

To see if the cloud bias implied by the incoming radiation comparison can be corroborated with independent data, we compare the HAR cloud cover climatology with MODIS remote sensing in Figure 3.11a. This does indeed confirm that the HAR underestimates overall cloudiness in the NWUIB, most severely in summer. Too little cloud therefore induces the summer underestimation of shortwave reflection identified in Section 3.4.4, which dominates total incoming radiation variability at this time of year. Similar conclusions regarding low cloud bias influences on shortwave radiation were also reached by Ruiz-Arias et al. (2016), albeit for a different climatic context. Interestingly, comparing the HAR cloud cover performance with the analysis in Forsythe et al. (2015) demonstrates that greater agreement with MODIS is evident for some global reanalyses, particularly ERA-Interim. Unravelling the causes of the apparent low cloud cover in the HAR is beyond the scope of this paper, but the microphysics, planetary boundary layer (PBL) and cumulus schemes might all play a role (e.g. Otkin and Greenwald 2008; Thompson et al. 2016; Ruiz-Arias et al. 2016).

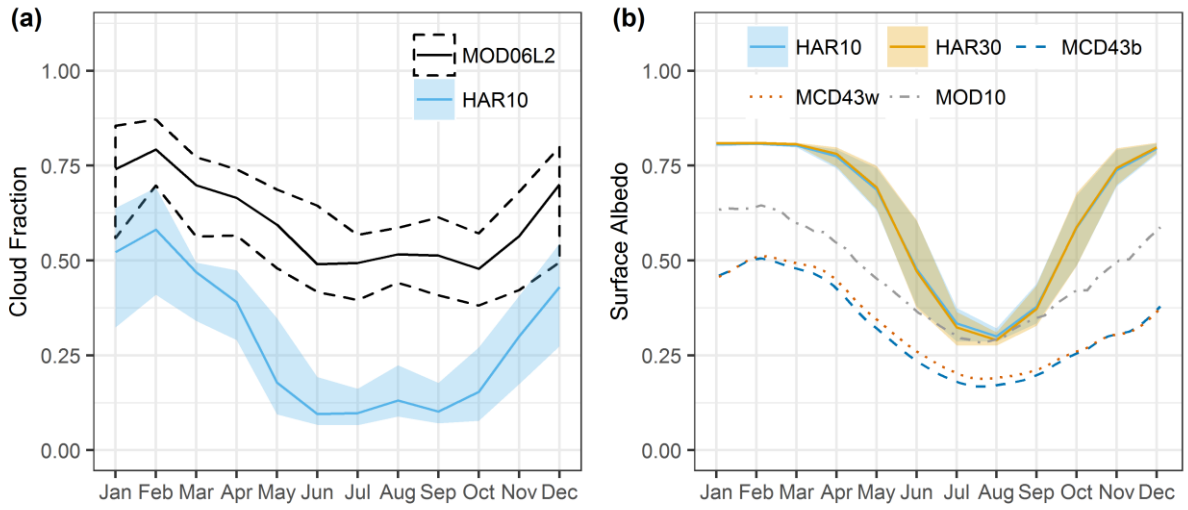


Figure 3.11 – Comparison of HAR cloud cover fraction (CCF) and albedo with MODIS products. The annual cycle of CCF for the NWUIB (spatial means) from MODIS (MOD06L2) is compared with the HAR in (a). Bold lines show the monthly means while ranges show the 10th to 90th percentiles (i.e. inter-annual variability). The HAR CCF variable is only available for HAR10. The annual cycle of spatial mean albedo for the NWUIB from MODIS and the HAR is shown in (b). MCD43b stands for MODIS MCD43A3 black sky albedo, MCD43w for MCD43A3 white sky albedo and MOD10 for

MOD10A1 snow albedo infilled using MCD43A3 for no-snow pixels. HAR ranges show the 10th to 90th percentiles (i.e. inter-annual variability).

Due to the important coupling between cloud radiative effects, snow cover and temperature (Betts et al. 2013, 2014; Forsythe et al. 2015), we also evaluate the HAR annual cycle of surface albedo, which is defined in general terms as the proportion of shortwave radiation received at the surface that is reflected. Figure 3.11b suggests that the HAR has notably higher albedo than MODIS in winter, spring and autumn. The largest overestimation occurs in spring, but there is potentially closer agreement in summer. The biases in albedo suggest that snow cover extent, the main control on the annual cycle of albedo variation in the UIB, does not decay sufficiently rapidly in spring and summer in the HAR. This issue is closely connected to the HAR's daily re-initialisation strategy, especially in terms of the challenges of deriving higher resolution snow cover (SWE) for each day from the much coarser FNL dataset. The snowpack process parameterisations in the Noah land surface model (LSM) may also play a role, as discussed in Section 3.5. While there is some uncertainty in the MODIS reference datasets (Section 3.3.3), the magnitude of differences in the mean, amplitude and shape of the annual cycles are large, which suggests that the HAR overestimation issues are substantive.

Relationships with Temperature

To examine the implications of the apparent biases in cloud and albedo, we focus now on correlations of monthly (ranked) anomalies of CCF and near-surface air temperature (Figure 3.12a), as well as correlations of surface albedo and temperature (Figure 3.12b). These correlations are performed separately for the HAR and observations (including MODIS) to see if the HAR adequately reproduces the observed dependencies between temperature and two of its controlling influences. From this we can infer whether there are process representation deficiencies in the HAR.

Considering observations first, Figure 3.12a shows a characteristic annual cycle of observed cloud influences on temperature. In winter, the positive correlations reflect the net warming effect of cloud cover at this time of year, which occurs through the mechanism of longwave enhancement (Forsythe et al. 2015). The subsequent transition to negative correlations shows a change in dominant cloud radiative effects to cooling in spring, summer and autumn. This cooling is induced primarily by cloud shortwave reflection, which increases in importance in line with the annual cycle of incoming shortwave radiation (Forsythe et al. 2015; Betts et al.

2013). At this time of year, Figure 3.12b shows that correlations between surface albedo and temperature in observations drop to around zero on average, although variability amongst station locations is notable. This is consistent with snow cover retreat in the spring and summer, which allows cloud cover variations to dominate net shortwave radiation variability and thus surface heating.

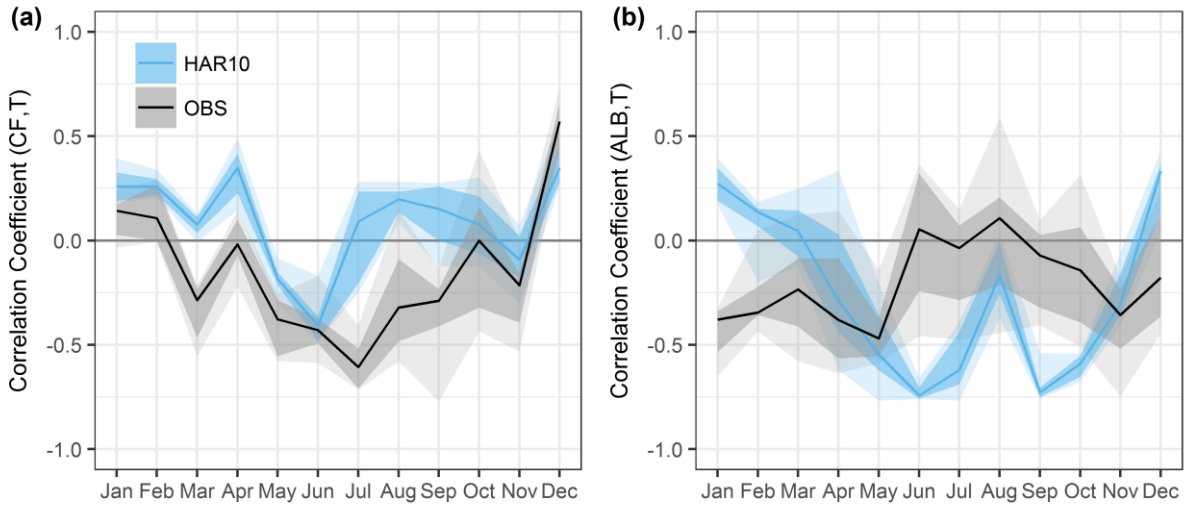


Figure 3.12 – Monthly correlations (Kendall's tau) between near-surface air temperature and selected variables. Correlations between temperature (T) and cloud cover fraction (CF) are given in (a), while correlations with albedo (ALB) are presented in (b). HAR variables are correlated with HAR temperatures, whereas observed correlations are based on in-situ observations of temperature in conjunction with MODIS albedo and cloud cover fraction corresponding to the station locations. Lines show the mean correlations and the successive shading shows the full and interquartile ranges based on the ensemble of stations.

Figure 3.12a suggests that the warming effects of cloud cover are present in winter, but the transition to negative correlations (cloud cooling effects) does not develop as early, strongly or consistently in the HAR. This is consistent with the concurrent low absolute cloud cover in the HAR, which suppresses the potential for inter-annual cloud variability to influence surface heating. For surface albedo, the HAR shows an annual cycle that is approximately the inverse of the pattern in observations. From spring into summer, with the exception of August, much stronger (negative) correlations between albedo and temperature exist in the HAR than in observations. The greater sensitivity to albedo variations likely reflects the influence of high biases in both the mean and variability of HAR albedo, but it is potentially compounded by the underestimation of cloud and its influences on temperature.

The implication of this is that the role of snow presence and its effects on albedo as a climate switch for cloud radiative effects is not entirely reproduced in the HAR (Betts et al. 2013, 2014). However, we note that the station locations analysed here are primarily in valleys, although Figure 3.11a does suggest that the cloud underestimation problem affects much of the NWUIB in summer.

The statistical significance of the correlations in Figure 3.12 is summarised in Appendix A (Section A7). This supports the point that there is a significant and physically coherent annual cycle of correlations between cloud and temperature in the observations, which does not develop as strongly in the HAR. It also confirms that the frequently significant sensitivity to albedo variation in the HAR in summer is higher than observations would suggest. However, we note that spread in Figure 3.12 can be large, which highlights the variability in these correlations and therefore the complexity of these relationships in reality.

3.5 Discussion

The results in Section 3.4 show both strengths and weaknesses in the HAR climatology for different variables. One important finding is that there are relationships between seasonal variations in bias in different variables. These relationships are relevant for various applications of the HAR dataset, as well as further regional climate modelling, such that we focus our discussion here on them.

Starting with temperature, the HAR bias is largest (coldest) in spring and smallest in summer. This helps to induce the annual cycle of bias in relative humidity, whereby a general high bias peaks in spring but reduces in summer. Critically, the seasonal variation in temperature bias appears to be associated with annual cycles of bias in the surface radiative balance. In summer, incoming shortwave (longwave) radiation is generally overestimated (underestimated), largely due to insufficient cloud cover. This would generally lead to greater surface heating and so reduce the overall cold (low) bias in temperature (DTR), albeit through error compensation. The relatively low fraction of annual precipitation falling in summer months is also consistent with insufficient cloudiness and reduced temperature bias. Furthermore, in spring, the peak cold bias in temperature and slow warming rates in the HAR appear to be related to the high bias in surface albedo. This is because surface warming, associated with rising incoming shortwave radiation at this time of year, would be impeded by overestimation of reflection effects.

The importance of albedo and snow representations for near-surface radiative and temperature biases was also recently investigated by Tomasi et al. (2017). Modifications to snow initialisation, snow cover fraction and albedo parameterisations in both the Noah and NoahMP LSMs were needed to reduce winter cold biases in their WRF simulations of an Alpine valley. For the Tibetan Plateau, Meng et al. (2018) additionally found that inserting albedo derived from MODIS into WRF simulations substantially diminished the cold bias in winter, spring and to a lesser degree autumn. However, García-Díez et al. (2015) noted that albedo may act mainly as a feedback amplifying poor representations of the snowpack or snow-atmosphere interactions in the Noah LSM, rather than a primary driver of temperature bias. Indeed, it may be that such a feedback underpins the large peak cold bias in spring in the HAR, possibly by reinforcing biases induced by other aspects of the WRF configuration. While snow process representation in Noah has improved over time (e.g. Livneh et al. 2010; Barlage et al. 2010; Wang et al. 2010), these and other studies (e.g. Saha et al. 2017) agree with our suggestion that some of the UIB cold bias in the HAR could potentially be reduced with revisions to the snow initialisation approach, as well as snow process and sub-grid variability parameterisations.

García-Díez et al. (2015) also highlighted the sensitivity of summer precipitation and temperature biases to different cumulus parameterisation schemes. Similar to our results, they demonstrated the potential for error compensation to reduce cold summer temperature biases through overestimation of incoming shortwave radiation. That summer humidity and temperature appear reasonably well simulated in the HAR, when cloud cover and precipitation are underestimated, confirms García-Díez et al.'s point that model evaluation needs to look at multiple climate variables to reveal possible error compensation and inconsistencies.

The other influences that may contribute substantially to biases in the HAR include PBL schemes, which parameterise unresolved turbulent fluxes of heat, momentum and moisture. Hu et al. (2010) showed that the differences in temperature when using various schemes relate to the degree of vertical mixing and entrainment of air overlying the PBL. Interestingly, Hu et al. found the Mellor-Yamada-Janjić (MYJ) PBL scheme used by the HAR to be the coldest of the schemes they tested, which could contribute to the general cold bias identified here. García-Díez et al. (2013) also found the MYJ scheme to be relatively cold, but they noted that the differences between PBL schemes were relatively consistent throughout the year, despite the fact that the sign of temperature bias changed between seasons. This led García-Díez et al.

to suggest that further investigation of the surface radiative balance is required to understand seasonal variation in temperature bias, as opposed to mean annual bias, which is supported by the results of this study.

Furthermore, several studies have shown that large errors can result from soil moisture initialisation using coarse resolution datasets (e.g. Case et al. 2008; Massey et al. 2016). Bastin et al. (2016) find that a dry soil bias leads to overestimated summer temperatures at their study site in France, with low cloud cover and consequent overestimation of incoming shortwave radiation acting to further dry the soil. The feedbacks between temperature, soil moisture, cloud cover and radiation biases in this example highlight the importance of interactions between multiple processes. Any local or regional soil moisture biases in the HAR could affect a range of surface and upper level variables through numerous feedbacks (Massey et al. 2016), particularly when snow cover declines in spring and summer. The scarcity of soil moisture data in the region for evaluation or assimilation is thus a significant issue, which is compounded by the challenge of providing sufficient spin-up with limited computational resources.

Several important avenues for future work arise from this discussion. One is that more sensitivity and performance tests with different physics parameterisation schemes are needed in this region. These would ideally be conducted with a multi-physics approach to test different microphysics, cumulus, PBL and LSM schemes (e.g. García-Díez et al. 2015). For the LSM scheme, the biases in the HAR suggest that testing needs to include careful evaluation of snowpack processes, snow-atmosphere interactions and sub-grid variability, as well as glacier representations. As noted by Collier et al. (2013), limitations arise in the Noah LSM used in the HAR from assumptions on minimum snow depth and SWE in glaciated cells, which likely overestimate albedo by omitting the influences of bare ice and debris cover under thin or no snow cover. MODIS snow cover and albedo products may help to further constrain the relevant process parameterisations to some degree.

In addition, the potential for improving initialisation strategies needs to be explored. This applies particularly to snow cover and surface albedo, which could again utilise MODIS remote sensing products (e.g. Meng et al. 2018). The NASA Soil Moisture Active Passive (SMAP) project may also ultimately provide a possibility to improve soil moisture initialisation, while the implications of spin-up period length should also be assessed. Moreover, it would be useful to test alternative data products for boundary conditions, such as

global reanalyses, in order to delineate the relative influences of WRF configuration choices and forcing datasets. In these tests, the boundaries of the inner model domain could be placed farther from the UIB, which would help determine if there is any residual boundary influence on the UIB in HAR10. With increasing computing power it may also be possible to test the feasibility of moving to higher resolution, convection-permitting climate simulations. This would be particularly useful for fine scale characterisation of near-surface climate, as well as hydrological and cryospheric modelling.

3.6 Conclusions

This evaluation of HAR performance for the UIB leads to three main conclusions. Firstly, we find that the HAR provides a good representation of UIB precipitation at multiple scales. In particular, HAR10 is consistent with most in-situ measurements, while also showing coherence with observed runoff for most gauged sub-basins. In addition, vertical precipitation gradients and spatial variation fit with other studies and inferences from MODIS data products. The HAR10 thus appears to represent a valuable source of precipitation data to supplement local observations in the UIB, although further testing through hydrological modelling should be undertaken.

Secondly, our results suggest that HAR temperature, humidity and incoming radiation in the UIB show reasonable climatologies overall, but also distinct patterns of seasonal variation in their biases. Temperatures exhibit a cold bias, which is largest in spring and smallest in summer. DTR is slightly underestimated throughout the year, with bias a little larger in spring and autumn. The HAR also displays high relative humidity in winter, spring and autumn, but less bias in summer. The available observations show that incoming shortwave radiation is overestimated in spring, summer (particularly) and autumn, during which time incoming longwave radiation is also underestimated.

Finally, we conclude that seasonal variation in biases is at least partially related to deficiencies in HAR cloud and albedo representations. Comparisons with MODIS products confirm that the HAR underestimates cloudiness, particularly during summer, which helps explain the biases in incoming radiation. Correlation analyses further suggest that the HAR does not fully reproduce the observed pattern of cloud radiative effects on temperature. Observations show a cycle of cloud warming effects in winter, through longwave enhancement, giving way to cloud cooling in summer, through shortwave reflection. This cycle does not develop to the same extent in the HAR. As such, the lower absolute

temperature bias in summer may stem at least partly from excess incoming shortwave radiation occurring in the HAR at a time when cloud reflection effects are strongly under-represented and surface albedo is at its lowest, thereby heightening warming. The high bias in HAR surface albedo throughout much of the year is also likely to affect the seasonality of biases, particularly in spring by reflecting a high proportion of the rising incoming shortwave radiation, magnifying the general cold bias.

Overall, our evaluation demonstrates the strong potential of simulations like the HAR for supplementing the limited local climate observations available in the Himalayan region. However, the findings also suggest that correctly parameterising cloud, snow cover and albedo processes appears to be critical for improved simulations of regional climate. Combining local observations and remote sensing data with more sensitivity tests may help to improve model representation, but of course uncertainties will be substantial in complex and data-sparse regions like the UIB. Further research is also required to see whether alternative datasets can provide improved boundary conditions and initialisation datasets for WRF.

Chapter 4

Feasibility and Skill of Process-Based Snow Modelling

4.1 Introduction

Snow processes in the Himalaya and adjacent ranges have profound influences on regional climate, hydrology and water resources (Bookhagen and Burbank 2010). As such, they need to be adequately represented in cryospheric, hydrological and land surface models. However, model applications in high mountains typically face two longstanding and inter-related problems (e.g. Klemeš 1990). Firstly, severe data paucity makes it difficult to drive and/or evaluate models. This is compounded by the substantial climatic variability that characterises mountain regions. Secondly, identifying appropriate model structures, parameterisations and parameter values is not straight-forward. This is especially the case in the Himalaya, where various processes operate and data to constrain modelling choices are particularly lacking.

For the data challenge, developments in high resolution regional climate modelling and remote sensing increasingly offer possibilities to supplement sparse local observations. For example, some studies have used the High Asia Refined Analysis (HAR) (Maussion et al. 2014) for offline forcing of glaciological and hydrological models (Huintjes et al. 2015; Biskop et al. 2016; Tarasova et al. 2016). However, studies like these are still few in number, covering only a small fraction of the Himalayan arc's diverse climate regimes. Crucially, there are especially few analyses characterising the skill of high resolution regional climate modelling to support snow model inter-comparison and identification. This also opens up the question of how far model (and input) evaluation in the Himalayan setting can be improved with remote sensing products that provide partial information on the surface energy balance, such as surface albedo and land surface temperature (LST) (e.g. Collier et al. 2013, 2015; Essery 2013).

These data advances help to partly confront the challenge of identifying appropriate model formulations, but this challenge is one that persists even in well-instrumented contexts. For example, in the SnowMIP and SnowMIP2 model inter-comparison studies, no single model formulation emerged as optimal, with performance varying between locations and years (Etchevers et al. 2004; Rutter et al. 2009; Essery et al. 2009). More recently, inter-comparisons using ensembles have generally also found there to be no single best model, but

instead groups of configurations that tend to perform consistently well, poorly or variably (Essery et al. 2013; Magnusson et al. 2015; Lafaysse et al. 2017). The systematic frameworks underpinning these latter analyses are an important advance for more precisely quantifying the influences of different parameterisations and identifying where improvements may be possible (Clark et al. 2015a). Furthermore, snow model complexity does not appear to be strongly (or necessarily positively) related to skill or transferability in space and time (see also Lute and Luce 2017), although data errors remain a large source of uncertainty even at well-instrumented sites.

While these site-based studies indicate which process representations may be generally robust, there have been very few analyses of transferability to the Himalaya. Indeed, there have been few applications of process-based, energy balance approaches here at all (Shrestha et al. 2015; Brown et al. 2014; but see e.g. Prasch et al. 2013). Yet these approaches are essential for testing hypotheses about snow processes, completing mass and energy balance frameworks in earth system models, and making projections in non-stationary climates. Therefore there is a need to better understand (1) how applicable process-based, energy balance snow models are, (2) how sensitive simulated snow processes are to modelling choices, (3) how transferable findings from model inter-comparisons in other regions are (i.e. performance), and (4) whether modelling improvements can be made using recent data advances.

As such, in this study we aim to evaluate the importance, skill and tendencies of different snowpack process representations in the Himalaya, focusing on (offline) hydrological applications in the Upper Indus Basin (UIB). Firstly, we assess the potential of relatively high resolution climate simulations to drive energy balance snow models. For this we use the HAR (Maussion et al. 2014), which has not been applied for modelling in the UIB so far. Secondly, we use the Factorial Snowpack Model (FSM) (Essery 2015) multi-physics framework to identify the importance, skill and tendencies of model configurations employing different parameterisations of key snowpack processes. We characterise ensemble performance, structure and variability by evaluating FSM process options at point and catchment scales, using local observations and multiple remote sensing products. Finally, we examine the significance of process parameterisation choices in an example application, focusing on runoff responses in a climate sensitivity test. Together, these aims contribute to the need for more application-oriented model evaluations using unified frameworks (Essery et al. 2013;

Clark et al. 2015a), in a real case where understanding snow hydrology is important for managing essential water resources.

4.2 Study Area

This study focuses primarily on the steep, mountainous Astore catchment (Figure 4.1). The Astore is a 3988 km^2 gauged sub-basin of the UIB, in which runoff is dominated by seasonal snowmelt (Archer 2003). This is derived primarily from orographically enhanced snowfall in the preceding winter and spring, which originates from westerly disturbances and strongly shapes inter-annual runoff variability (Archer and Fowler 2004). Snow redistribution processes are important but poorly quantified, whereas glacier cover is relatively limited, at around 6% according to the Randolph Glacier Inventory (RGI) 5.0 (Arendt et al. 2015). In addition, the ESA GlobCover 2009 product (Arino et al. 2012) confirms that vegetation cover is relatively sparse, with the catchment dominated by a mixture of bare ground, herbaceous plants, and perennial snow and ice.

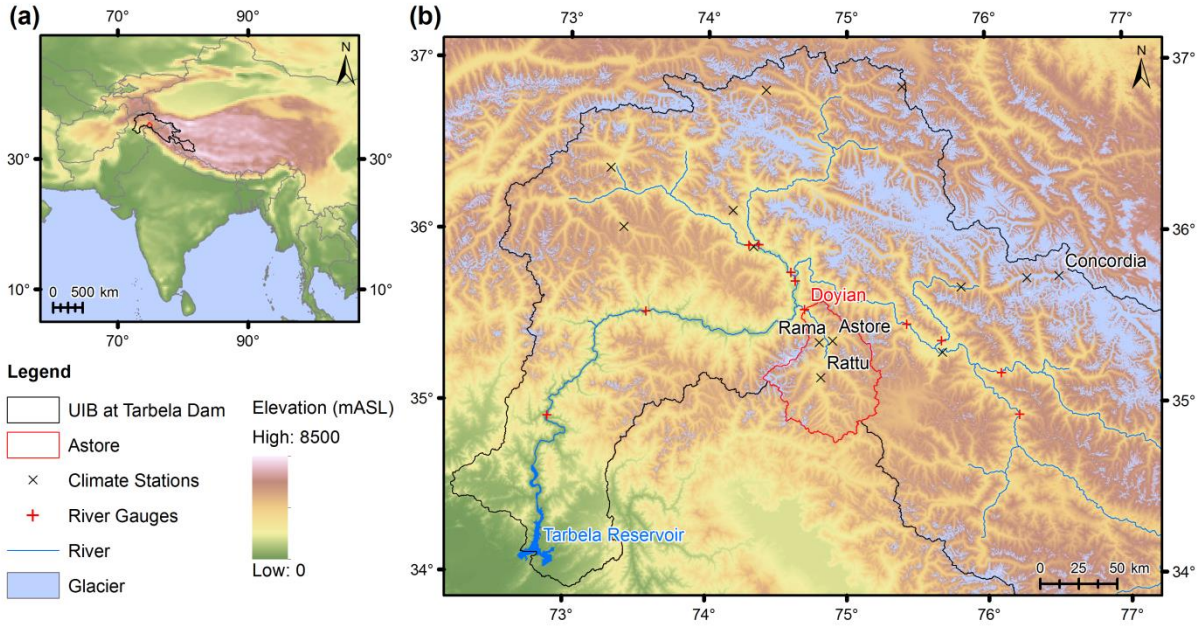


Figure 4.1 – Location of study area and local measurement points. The regional context is indicated in (a). The Astore catchment and observation locations (with labels for the most important sites in this study) are shown with topography and glacier extent in (b). This includes the Concordia site to the east of the Astore catchment, which is used for model validation at the point scale, as discussed in Section 4.3.5.

4.3 Data and Methods

The methods applied here to evaluate the skill and tendencies of alternative snow model configurations in the Himalaya may be summarised as follows. Firstly, we adapted the FSM model to additionally permit catchment scale simulations. Secondly, we applied minimal bias correction and simple downscaling methods to the HAR in order to derive distributed input fields. Then we used local observations and multiple remote sensing products to help evaluate the model ensemble. Finally, to briefly assess variations in climate sensitivity within the ensemble, we applied simple perturbations focusing on temperature. In the following sections we elaborate this approach.

4.3.1 Model

FSM Overview

The FSM multi-physics ensemble employed here is an intermediate complexity, systematic framework for testing alternative representations of key snowpack processes and how they interact with each other within a coupled mass and energy balance scheme (Essery 2015). With two parameterisation options (0/1) for five different snowpack processes, the FSM ensemble includes 32 possible model configurations. The five snowpack processes and their parameterisation options, which neatly synthesise common approaches in a range of widely applied models, are summarised in Table 4.1. Analyses of FSM to date have shown that it gives ensemble performance and spread comparable to larger multi-model ensembles (Essery 2015). Its value for testing new process representations has also been demonstrated (Moeser et al. 2016).

Adaptations and Implementation

This study adapts the initial point-scale FSM implementation for distributed simulations on a regular grid, with spatially varying inputs. There are several reasons why this is useful. Crucially, appropriate data and sites for point-scale simulations in the Himalaya are very scarce and of limited spatial representativeness. Hydrological (and other) applications need to consider catchment heterogeneity, including the large elevation, aspect, slope and related near-surface climate variations that are critical to hydrological functioning in basins like the Astore. The core subroutines of FSM are left as described in Essery (2015), and each horizontal grid cell is simulated independently of other cells, i.e. inter-cell mass and energy

transfers are not considered, as in other distributed snow models (Liston and Elder 2006a; e.g. Lehning et al. 2006).

Process Description	Short Name	Parameterisation 0	Parameterisation 1
Snow albedo variation	Albedo (AL)	Diagnostic – function of surface temperature	Prognostic – decays with time and increases with snowfall
Density of fresh snow and snowpack density evolution	Density (DE)	Constant	Specified fresh snow density and compaction increases with time
Liquid water storage, drainage and refreezing	Drainage (DR)	Instant drainage, no refreezing	Bucket model (drainage to layer below if liquid holding capacity exceeded), with refreezing (and latent heat release) allowed
Atmospheric stability adjustment for turbulent heat fluxes	Stability (S)	No adjustment for atmospheric stability	Stability factor is a function of the bulk Richardson number
Thermal conductivity for heat conduction	Thermal Conductivity (TC)	Constant	Function of density

Table 4.1 – Summary of the process parameterisation options available in FSM. Full details are given in (Essery 2015). The short names by which the processes are referred to in the text are given.

The simulations use a 500 m horizontal resolution grid and an hourly time step, with topography from the SRTM 90 m DEM v4.1 (Jarvis et al. 2008). Spatial variation in land surface properties is ignored on the basis that glacier and vegetation (including forest) cover are relatively low, while information on substrate properties is unavailable. Subgrid variability is not incorporated or (further) parameterised in these reasonably high resolution simulations, and snow redistribution by avalanching and wind are omitted. While some limitations on absolute model performance could follow from omitting these complicated factors, our emphasis is more on relative performance and the differences arising from alternative representations of snowpack processes. For this reason, we also opt not to apply hydrological routing in this steep, fast-draining catchment. The properties and geometries of both substrates and channels are unknown and typically calibrated even in process-based models (e.g. Ragettli et al. 2013; Shrestha et al. 2015), which could confound interpretation of

snowpack processes here. However, future work introducing these additional processes and complexities should be undertaken.

4.3.2 Climate Inputs

High Asia Refined Analysis (HAR)

The HAR (Maussion et al. 2014) is used as the basis for spatiotemporally varying input fields of rainfall, snowfall, air temperature, relative humidity, wind speed, surface air pressure, and incoming shortwave and longwave radiation. The HAR is effectively a 14-year dynamical downscaling of coarser global analysis to 10 km over the Himalaya and Tibetan Plateau, using the Weather Research and Forecasting model (WRF, Skamarock et al. 2008). The skill of the HAR in capturing aspects of regional near-surface climatology was demonstrated by Maussion et al. (2014). The evaluation in Chapter 3 used local observations and remote sensing to analyse HAR performance in more depth for the UIB. While some biases are present, the HAR has been shown to exhibit a skilful representation of climate in several hydrological and glaciological modelling studies in other parts of the region (Huintjes et al. 2015; Biskop et al. 2016; Tarasova et al. 2016).

Bias Correction and Downscaling

Following preliminary testing of different strategies, a minimal bias correction approach was adopted. Specifically, we corrected near-surface air temperatures for biases in the mean, based on local observations in the Astore catchment (Figure 4.1b). We do not apply more detailed or complicated correction methods for several reasons. Firstly, the HAR shows many areas of good performance across a range of variables. Secondly, it is difficult to fully characterise spatial and temporal variation in biases based on limited observations, especially for less commonly observed variables. Indeed, after testing alternative approaches, including quantile-based methods using observations from neighbouring catchments, it was found that the methods attempted did not lead to improvements in cross-validation or model performance. Finally, we aim to retain the physical consistency of the HAR fields as far as possible, in terms of both inter-variable and spatiotemporal dependence structures. Clearly substantial uncertainty exists for the resulting input fields, which we address below.

The downscaling method draws on the approaches within the MicroMet meteorological pre-processor of SnowModel (Liston and Elder 2006b), as well as the methods used by

Duethmann et al. (2013). With the exception of incoming shortwave radiation (see below), the method here uses linear regression to relate each variable to elevation, a primary control on climate variation here, with prior log-transformation for precipitation. The values at each HAR cell are then interpolated to a reference level using the gradient determined in regression. This gives a spatial (horizontal) anomaly field. This field is interpolated to high resolution, and then the elevation signal is reintroduced using the regression gradient. There are of course other possible methods and refinements that could help to bridge the scale gap, but analysing these is not the focus of this study. It is of course difficult to fully determine the effectiveness of this or any downscaling method from limited observations, but additional confidence can be gained from model ensemble validation (Section 4.4.1).

Given the pronounced topography of the study area, we model clear-sky shortwave radiation at the surface based on the high resolution DEM using the approach developed by Corripio (2003). This approach has been successfully applied before in this region (e.g. Ragettli et al. 2013) and additionally checked against measurements from EvK2CNR stations. We include the effects of variations in cloud/atmospheric transmissivity by calculating the ratios of clear-sky to received incoming shortwave radiation at the surface in the HAR. We interpolate this to the higher resolution of the model grid and then adjust the clear-sky values accordingly. This approach helps to maintain consistency between variables and cloud influences while also capturing key topographic effects, although direct/diffuse partitioning and spatial heterogeneity of shortwave inputs under different cloud conditions are clearly simplified.

Uncertainty

In recognition of input uncertainty, we also perform simulations using two alternative input strategies. These are summarised in Table 4.2 below. The strategies are not independent, as their main purpose is to indicate whether the conclusions reached on snowpack process representations, the focus of this study, are unduly affected by the approaches described above. This is of course only a partial treatment of uncertainty. However, the baseline analysis presented here is a precursor to more detailed investigation of the difficult problem of how to quantify uncertainty in climate fields when observations are so sparse.

Input/Run Name	Input Strategy
FSM1	Downscaled HAR using approach described in Section 4.3.2, with simple bias correction of temperature fields
FSM2	Downscaled HAR using approach described in Section 4.3.2, but no bias correction of temperature fields
FSM3	Downscaled HAR precipitation as per Section 4.3.2, but other fields primarily estimated from observations. Specifically, temperature is lapsed based on observations (separately for daily minima and maxima, using monthly lapse rates), and the month's climatological hourly diurnal cycle from EvK2CNR stations is applied. Relative humidity is estimated from daily minimum and maximum observations, and disaggregated using a similar approach. Incoming shortwave radiation is calculated as per FSM1 and FSM2, but a parameterisation of cloud transmissivity based on diurnal temperature range is used (Pellicciotti et al. 2011). Incoming longwave radiation is estimated as in MicroMet (Liston and Elder 2006b). Wind speed is taken from the HAR in the absence of observations.

Table 4.2 – Model input strategies reported in this study.

4.3.3 Model Evaluation

In line with the study's hydrological emphasis, the overall volume and timing of simulated runoff is evaluated using daily flows recorded at the Doyian gauging station by the Water and Power Development Authority (WAPDA). No local snow measurements are available, but we compare model outputs with SCA derived from the MODIS MOD10A1 product (Collection 6) (Hall and Riggs 2016), as a further spatial constraint on catchment internal processes. We also explore the potential for MODIS albedo (MCD43A3 and MOD10A1) and LST (MOD11A1) products to support evaluation of processes affecting the surface energy balance. These products have been widely evaluated (Liu et al. 2009; e.g. Wan et al. 2004), although additional challenges are posed by complex terrain. Additional evaluation of LST is given in Appendix B (Section B1), which confirms that the MODIS product generally performs well, with low enough bias to indicate gross model errors. Each data source has its own limitations and uncertainties of course, but in combination they permit the qualitative and quantitative performance evaluations reported in Section 4.4.

4.3.4 Climate Sensitivity

To contextualise findings on how snowpack process representation choices affect model skill and tendencies, we perform a simple climate sensitivity test. For this we perturb summer air

temperatures (April to September) by -3 to +3°C, in 1°C increments. The approach taken here is essentially that used in other climate sensitivity studies (e.g. Rasouli et al. 2014; López-Moreno et al. 2017), but without examining precipitation perturbations. In fact we keep October to March precipitation and temperature deliberately as per the baseline, in order to focus on differences arising in the melt season. The approach is clearly very simple and does not fully account for the spatiotemporal (and inter-variable) dependencies and variation that characterise more realistic climate sensitivity and change scenarios. However, it should be sufficient to see whether model configuration choices elicit notable differences in hydrological response, an important consideration for climate change impact studies.

4.3.5 Point-Scale Simulations

While our analysis here focuses on distributed simulations, we also performed point-scale runs with FSM for the Concordia monitoring site set up and run by EvK2CNR (Figure 4.1b). This site measures all input variables needed for FSM, but data are only available for a couple of years. The purpose of the point-scale simulations here is to check the performance of the FSM code and to provide selected additional checks on the findings of the distributed simulations. Full details are given in Appendix B (Section B2) and in Section 4.4 where relevant.

4.4 Results and Discussion

This section begins by examining overall ensemble performance for runoff and SCA (Section 4.4.1). The significance, tendencies and realism of different process parameterisation choices are then evaluated and related to ensemble structure and variability, which helps to identify which configurations are likely to be most robust (Sections 4.4.2-4.4.4). To contextualise the importance of model parameterisation choices, a brief assessment of sensitivities to climate perturbations then follows (Section 4.4.6).

4.4.1 Ensemble Evaluation

Runoff

Figure 4.2a shows that the modelled mean annual cumulative runoff curves are in overall agreement with observations, in terms of both shape and magnitude. However, the spread in the FSM ensemble is notable. This takes the form of groupings of ensemble members. Three

clear groups are distinguishable early in the melt season, but these groups subsequently split to give a more complicated ensemble structure. Focusing on the earlier part of the season, when snowmelt is expected to dominate runoff composition, the closest correspondence with observations is for the group with the shallowest runoff curve. The other groups tend to show significantly faster runoff rises. The differences between groups are significant from hydrological and water resources perspectives, with cumulative runoff by the beginning of June (July) around 2 (1.5) times larger in the fastest-responding group compared with the slowest.

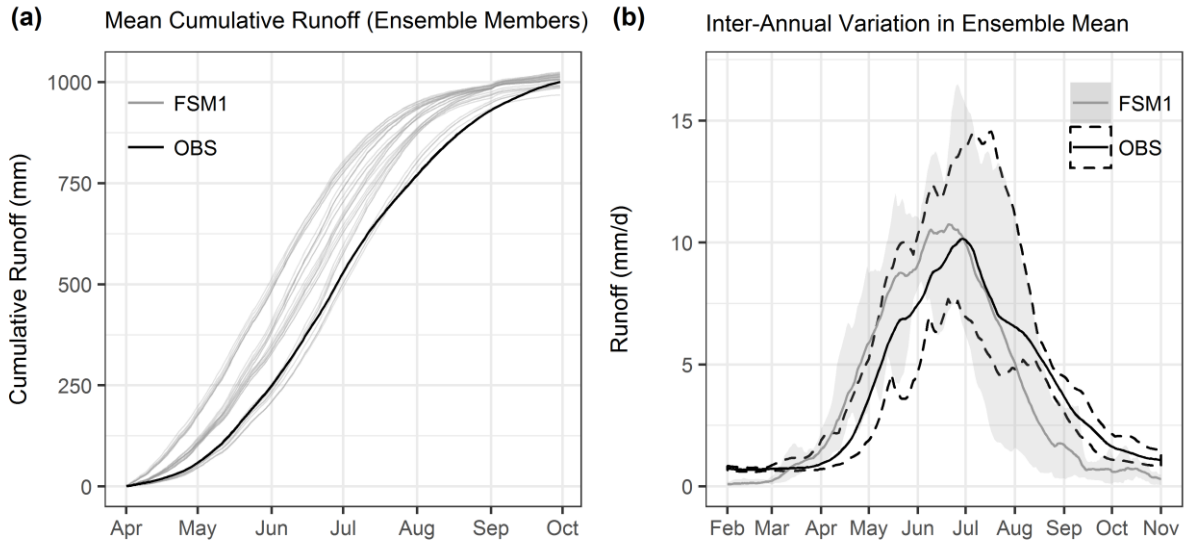


Figure 4.2 – Comparison of modelled snowmelt runoff with observed runoff. Mean annual cumulative runoff for the high-flow season for each of the 32 ensemble members is given in (a), along with observations. Inter-annual variation in the ensemble mean and observed runoff is shown in (b), where the solid lines denote the mean and the shaded/dashed areas indicate the inter-annual range. Model results are based on the FSM1 inputs and are shown in grey, with observations in black.

To assess inter-annual variability, we compare the annual runoff cycle for the ensemble mean with observations in Figure 4.2b. The main point from this is that the modelled inter-annual mean and range of runoff are consistent with observations. Clearly the timing of melt onset is a little earlier than observed runoff for the ensemble as a whole, as expected from Figure 4.2a, while the modelled hydrograph also exhibits a steeper and more variable falling limb. These features may be partly related to the modelling simplifications discussed in Section 4.3.1, especially omission of routing and glaciers, but particularly to the fact that the ensemble mean might not provide the best representation of melt and runoff processes. Individual members or certain configurations may be better, as indicated in Figure 4.2a and explored in detail below

(Section 4.4.4). If the timing difference is then temporarily ignored, good correspondence in the range of inter-annual variability is apparent.

Snow Cover

Figure 4.3a demonstrates that the mean annual cycle of catchment snow cover variation is reproduced by the FSM ensemble. As with runoff, notable spread between ensemble members is present, especially in spring and early summer. During this period, most groups of ensemble members tend to lie between the two MODIS Normalised Difference Snow Index (NDSI) thresholds used here for snow cover classification (see figure caption). That most members are below the NDSI 0 threshold cycle (i.e. classification based on very limited or no snow cover in a pixel) is indicative of a faster decline (and slower subsequent rise) of modelled SCA compared with MODIS for the ensemble as a whole. This appears consistent with the finding above that modelled snowmelt runoff begins and rises earlier than in observations, potentially due to overly large spring melt rates. However, input uncertainty, scatter in the relationship between NDSI and fractional snow cover (Salomonson and Appel 2004), and modelling simplifications are amongst the important caveats here. We return to this in more detail below.

From Figure 4.3b it can be seen that the inter-annual range of SCA throughout its annual cycle is broadly consistent with MODIS. While both MODIS series are variably offset from the modelled ensemble mean cycle, their ranges are generally of similar magnitude in all seasons. Variation in both MODIS and the model is small during the period of extensive snow cover in winter, before increasing during spring and summer. This melt season variability reflects important inter-annual and spatiotemporal variations in both mass and energy inputs. The largest inter-annual variation in snow cover occurs during autumn, reflecting variability in the timing of significant snowfall events at this time of year.

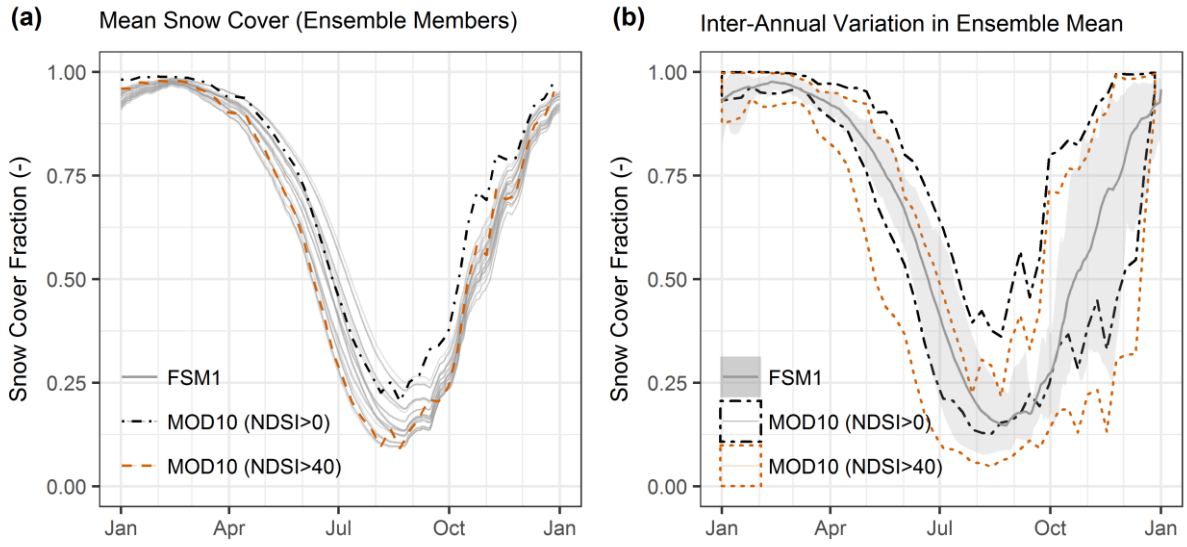


Figure 4.3 – Comparison of modelled catchment snow cover with MODIS remote sensing. The mean annual cycle of snow cover for each ensemble member is displayed in (a), along with MOD10A1 snow cover. Inter-annual variations in the ensemble mean and MOD10A1 snow cover is plotted in (b) as the dashed/shaded areas. Model results are based on the FSM1 inputs. The two MOD10A1 series are in black and orange, which correspond to Normalised Difference Snow Index (NDSI) thresholds of 0/100 and 40/100. Two thresholds are used to show the differences between how classifying pixels using approximately 0% and 50% snow cover affects SCA. The former threshold is more consistent with the model results (in grey), in which grid cells are classified based on whether or not any snow is present.

In conjunction with the runoff evaluation above, this comparison therefore suggests that ensemble performance is adequate for further investigation of the importance, skill and tendencies of different snowpack process representations. This enables us to understand how they shape ensemble spread and grouping structures, and ultimately to quantify performance variations. We note that the simulations using the two alternative input strategies (Section 4.3.2) also show reasonable runoff and SCA performance (Appendix B, Section B3), such that they are used for comparison in subsequent sections. The lowest performance is seen in the more observation-based approach (FSM3), which also shows the most spread in both the ensemble and its groupings.

4.4.2 Process Importance

Figure 4.4 helps us to gain an initial idea of the relative importance of the processes options in FSM for ensemble spread and structure in runoff. For each of the five snowpack processes, we average the ensemble members that use the 0 and 1 parameterisation options and then calculate the differences between these averages. From doing this we can see that the albedo,

drainage and stability processes are associated with notable deviations from zero. This means that they have important effects on ensemble spread in melt and runoff, similar to initial results from the French Alps (Essery 2015). Notably, the drainage process shows opposite signs of difference for melt and runoff initially. Inter-annual variability in differences is also substantial, particularly for the albedo option. The patterns of differences and their causes are explored in detail below, but for now we note that it is these three processes that are most critical for hydrological responses.

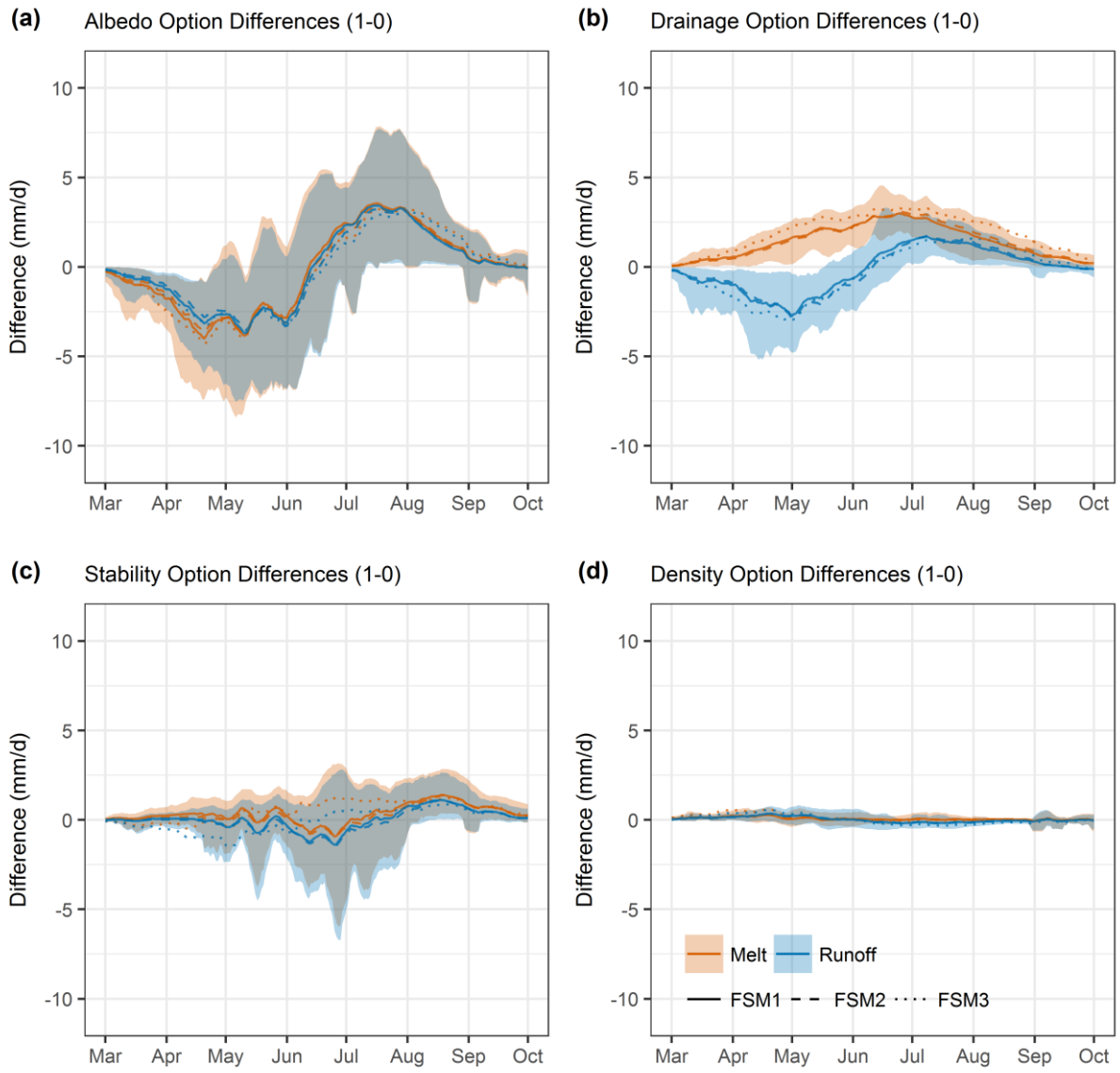


Figure 4.4 – Differences in simulated daily melt and runoff according to model options. For a given process, differences are calculated each day by subtracting the mean of all ensemble members using option 0 from the mean of all ensemble members using option 1. The differences due to albedo (a), drainage (b), stability (c) and density (d) options are plotted. Differences in melt and runoff are shown in orange and blue respectively. Line types denote the inter-annual mean for each input strategy, while the shading shows inter-annual ranges for FSM1 only for clarity.

In contrast, the density process options show small differences in Figure 4.4d, which is also the case for thermal conductivity (shown in Appendix B, Section B4). However, this does not necessarily mean that these processes are unimportant, either for hydrology or other variables of interest. This is further explored in Appendix B (Section B4), where we show that the density option exerts influences on runoff timing in some years, especially later in the melt season. While we focus mainly on albedo, drainage and stability adjustment options in subsequent sections, as the primary controls identified in Figure 4.4, the hydrological implications of density and thermal conductivity options should be explored further.

4.4.3 Process Evaluations

Having established the relative importance of different processes, the hydrological tendencies arising from each of the albedo, drainage and stability adjustment options are now explored in turn and evaluated with additional data sources as far as possible. This is followed by an assessment of spatiotemporal variation.

Albedo

In Figure 4.4a we saw that the sensitivity of model response to the albedo option is large, in line with the known importance of net shortwave radiation as a driver of melt in this environment (Hewitt 2014). Specifically, switching on the prognostic albedo option leads to lower melt and runoff in the first part of the melt season, a pattern which reverses around the middle of the season. The prognostic albedo parameterisation thus initially delays and slows melt relative to the diagnostic option, with one implication being that higher SWE persists longer into the melt season when melt energy (and melt rates) are higher. While the mean differences are similar for each input strategy tested, inter-annual variability is substantial.

The main reason for the faster melt in spring and early summer using the diagnostic parameterisation based on LST is its pronounced diurnal cycle. This is demonstrated in Figure 4.5a, which shows that the diagnostic parameterisation has substantially lower albedo during the afternoon hours. As such, there is a feedback here, whereby rising incoming shortwave radiation increases LST, in turn decreasing simulated albedo and leading to higher net shortwave receipt. The prognostic parameterisation does not induce such a pronounced diurnal cycle, due to the longer timescales of albedo decay it represents. This explains why faster melting occurs with the diagnostic parameterisation in the earlier part of the melt season, when snow is abundant.

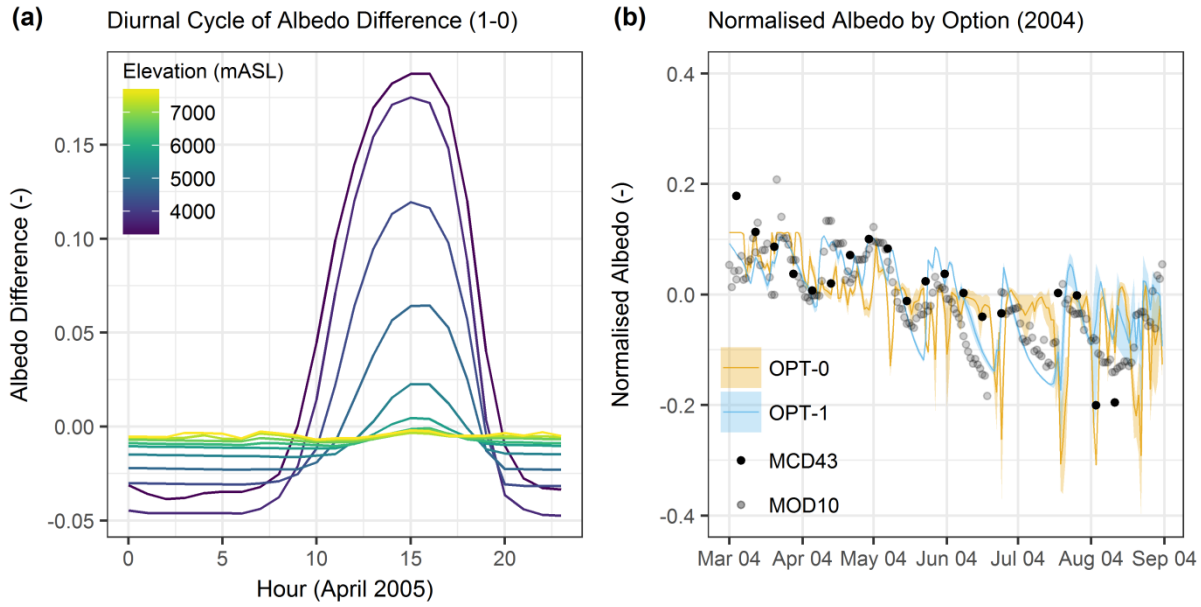


Figure 4.5 – Comparison of albedo parameterisations. The average diurnal cycle of differences in snowpack albedo arising from the two parameterisation options is shown in (a), for an example month and for different elevation bands (coloured). Differences are based on the averages of ensemble members using each option. The differences are calculated as prognostic (1) minus diagnostic (0), such that positive values indicate higher albedo with the prognostic parameterisation. In (b) we compare modelled catchment-average (normalised) snow albedo with MODIS remote sensing for an example year (2004). Modelled albedo is grouped by option, with diagnostic (0) in orange and prognostic (1) in blue. Both the mean (line) and range (shading) of ensemble members are shown. MCD43A3 is shown with black dots, while MOD10A1 (8-day moving averages) is shown with grey dots. The modelled series are normalised by subtracting the ensemble mean albedo (all members), while the MODIS series are normalised by subtracting their respective means. Normalisation was undertaken to emphasise functional behaviour rather than differences in absolute albedo, which are clearly important but more uncertain due to the challenges of albedo retrieval in complex terrain and the simplified treatment of subgrid variability in the model.

While albedo does vary diurnally with solar zenith angle in reality, it does not necessarily follow that this is related simply to LST cycles or that the diagnostic approach is more accurate over daily and longer timescales. To independently evaluate the realism of the two albedo parameterisation options, we compare daily time series of modelled catchment-average snow albedo with MODIS remote sensing in Figure 4.5b. Acknowledging some timing offsets, this comparison suggests that the functional form of the prognostic parameterisation is much more consistent with remote sensing. In particular, this option more skilfully captures the sharp albedo increases following snowfall, which are an important factor behind melt rate variability in the spring and summer (Hewitt 2014). Moreover, the prognostic option better reproduces the shape of albedo decay. This can reduce albedo below the values from the

diagnostic parameterisation later in the season, which also contributes to the higher melt rates at this time when using the prognostic parameterisation.

Point-scale simulations at Concordia (Section 4.3.5) further confirm that the prognostic option is more realistic (Appendix B, Section B5). Moreover, in data from Concordia and MODIS, the relationship between albedo and LST shows substantial scatter (Appendix B, Section B5). This further suggests that calibration of parameters in a simple diagnostic approach would struggle to rectify shortcomings in its functional form. The prognostic approach should therefore be preferred in hydrological models and LSMs.

Drainage

In Figure 4.4b we saw that melt is higher with the drainage option switched on. Conversely, runoff is initially lower, before switching midway through the melt season. This pattern results from the interactions of refreezing and storage tendencies in the snowpack. With the drainage option on, liquid water from melting is allowed to refreeze, leading to latent heat release, which maintains a higher snowpack temperature (confirmed in Appendix B, Section B6). This is conducive to higher melt rates, but it does not lead to higher runoff rates initially. Retention and delayed release of liquid water using the bucket approach are part of the reason for this, but importantly, multiple diurnal cycles of melting and refreezing may be required before a given unit of snow is entirely converted to runoff. The delaying effect of switching on the drainage option thus outweighs its tendency to increase melt rates. Yet by allowing snow to persist for longer, this enhanced storage ultimately leads to higher melt and runoff rates later in the season, as later-lying snow becomes subject to increasing energy inputs. Unfortunately data to evaluate these processes directly are unavailable, but we discuss the implications of these tendencies further below.

Stability Adjustment

Although inter-annual variability in Figure 4.4c is large, the mean difference series suggest that switching the stability adjustment option on leads to slightly lower melt and runoff rates in the middle of the season, but then higher rates subsequently. The primary explanation for this lies in the snowpack surface energy balance. Figure 4.6a shows that, during the early part of the year, the differences in turbulent fluxes arising from the stability adjustment choice are largely offset by differences in net radiation. However, as air temperatures and near-surface gradients increase in spring and summer, the net turbulent flux difference ultimately becomes

a key driver of the differences in the surface energy balance residual (i.e. the additional energy available for warming/melting the snowpack when no stability adjustment is applied).

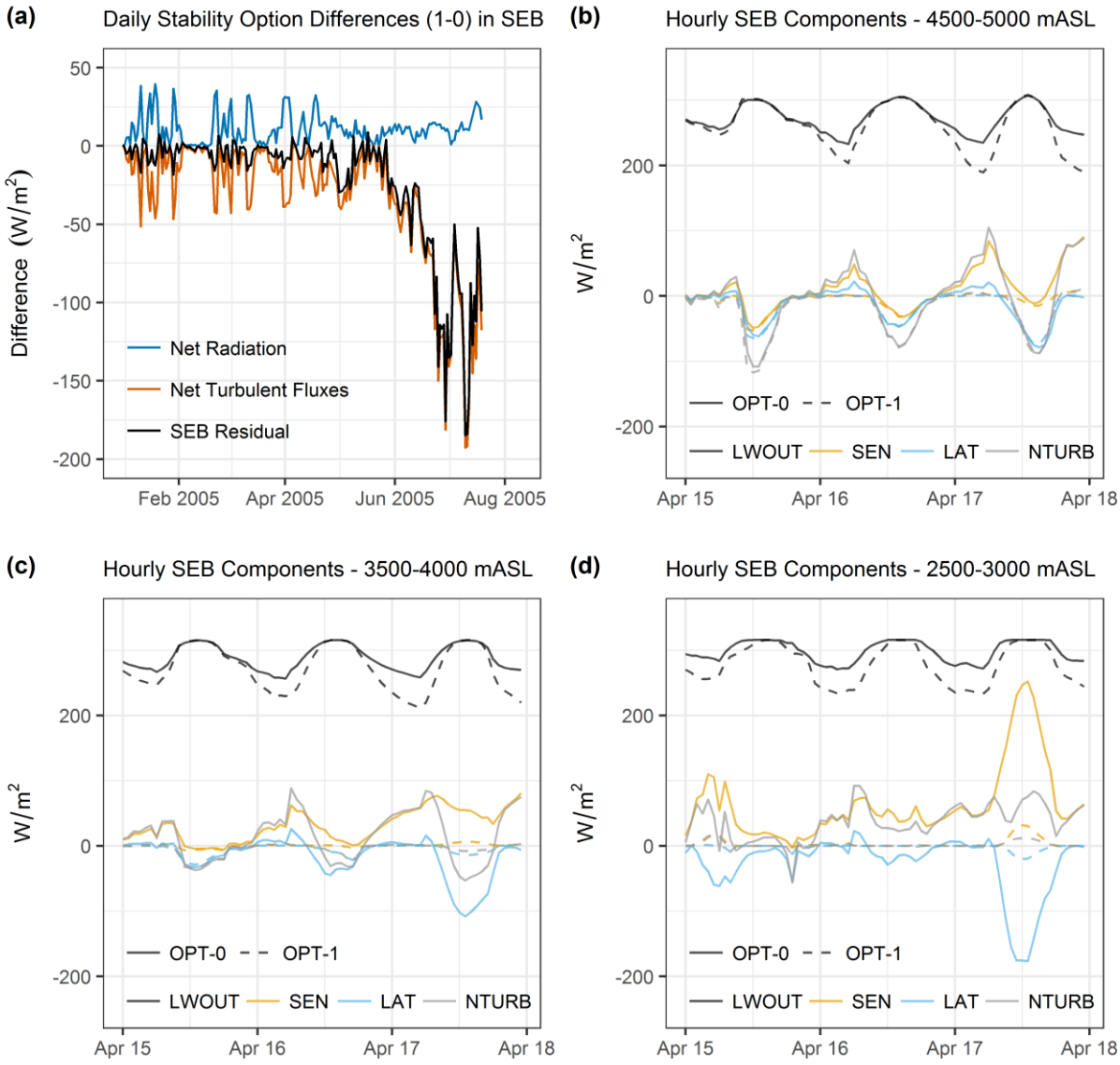


Figure 4.6 – Stability adjustment influences on snowpack surface energy balance (SEB) components. Daily time series of differences in snowpack SEB composition arising from stability option choice are shown in (a) for an example elevation band (3500-4000 mASL) transitioning from winter through to melting in summer. In (b) to (d) we show hourly time series for 3 days in April 2005 for successively lower elevations, which essentially represent non-melting, early melting and intense melting conditions. These panels show the averages of outgoing longwave radiation (LWOUT), as well as sensible (SEN), latent (LAT) and net (NTURB) turbulent heat fluxes, which are averaged for ensemble members using the 0 (no adjustment – solid line) and 1 (adjustment – dashed line) stability options. Results are based on FSM1 inputs.

This progression is shown through example hourly time series in Figure 4.6b-d. These series correspond with non-melting, transitional and melting conditions. Specifically, from Figure

4.6b we can see that, when no stability adjustment is applied, outgoing longwave radiation increases to largely offset the additional turbulent heat transfer towards the surface, which is driven by the sensible heat component. This compensation occurs principally at night, with turbulent fluxes converging during the day. Figure 4.6c typifies the emergence of notable sensible heat fluxes towards the surface during the day as temperatures warm, which is most clearly developed in the late-lying snow case in Figure 4.6d.

As turbulent flux measurements are unavailable, we examine whether LST from MODIS remote sensing can provide any partial constraints on the flux representations and their contribution to the surface energy balance (see Section 4.3.3 and Appendix B Section B1 for additional validation of MOD11A1 product). Figure 4.7 shows that the largest differences in vertical LST profiles due to stability option choice occur at night and increase with elevation, for the clear-sky conditions when MODIS retrievals are available. Switching off the stability adjustment generally leads to profiles that are significantly warmer and more consistent with MODIS, due to efficient near-surface mixing. In contrast, ensemble spread in day-time LST is relatively small and generally in good agreement with MODIS. There are slightly larger differences in summer, although the extent and influence of sub-pixel snow cover variation on MODIS LST likely increases during melting periods, giving some positive bias (Appendix B, Section B1).

As air temperatures and LST generally converge here at night-time (Appendix B, Section B1), the simulated differences at night may stem from too strong a suppression of turbulent fluxes using the bulk Richardson number to correct for atmospheric stability. Importantly, the identification of this at night-time suggests that a similar issue could be present during the day under snowmelt conditions. There could also be other complications here, including (poorly observed) incoming longwave radiation and terrain enhancement effects (Sicart et al. 2006), as well as subgrid variability (Clark et al. 2011) and sensible and latent heat advection (e.g. Harder et al. 2017). The results here suggest that testing the inclusion of these effects, as well as using additional approaches for stability adjustment (see e.g. Andreadis et al. 2009), is therefore warranted in this context. The potential for error compensation is high, but the ensemble framework approach of FSM could allow for exploration of predictive uncertainty at the very least.

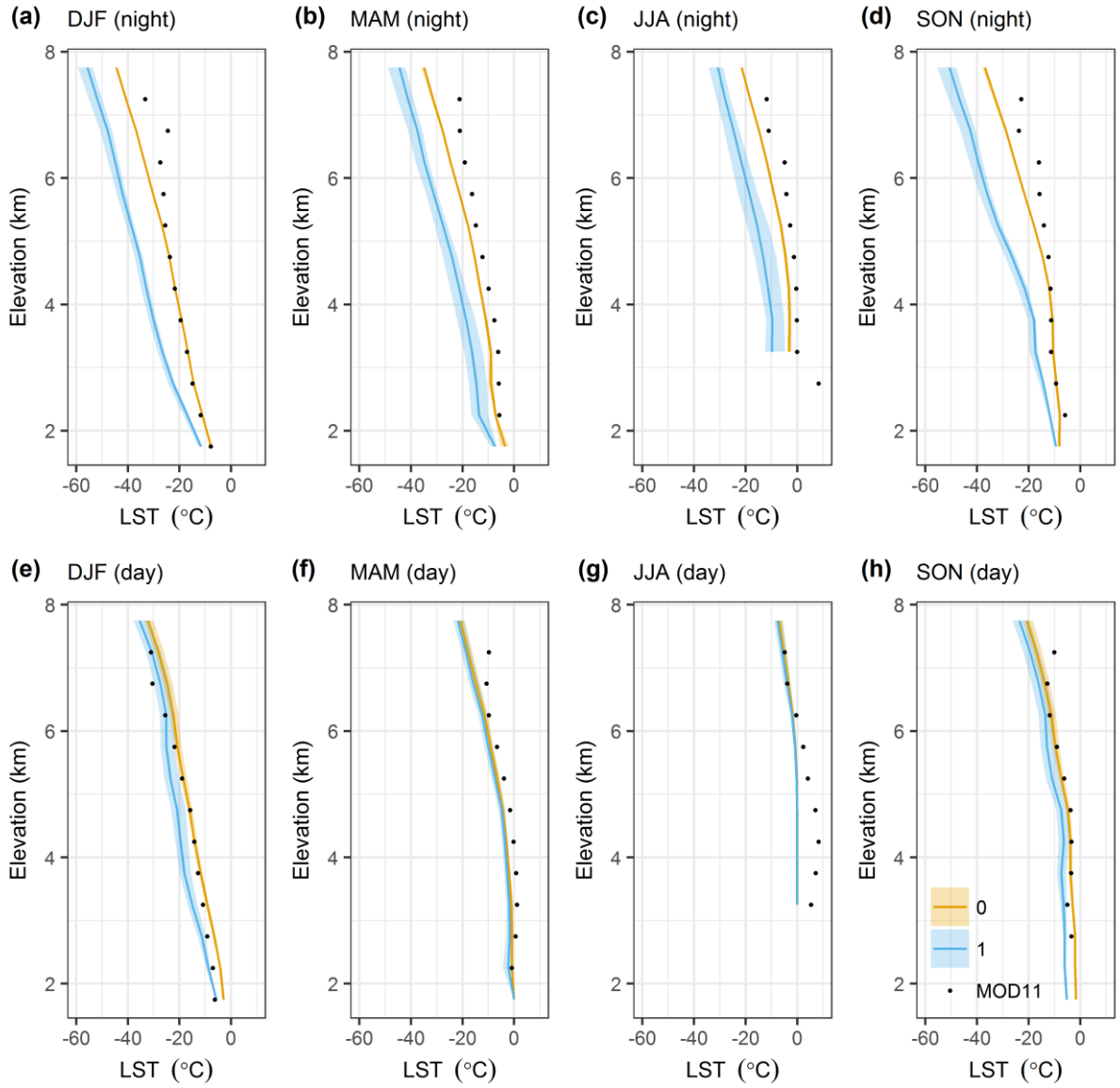


Figure 4.7 – Comparison of modelled seasonal mean elevation profiles of LST (stratified by stability option) with MODIS MOD11A1 remote sensing. In (a) to (d) we show night-time temperatures, with day-time temperatures in (e) to (h). Model results are based on the FSM1 inputs and correspond with the closest model time step to the Terra platform overpass times, as well as only days for which MODIS retrievals are available (i.e. clear-sky conditions). The 0 (no adjustment) and 1 (adjustment) stability options are shown in orange and blue, respectively, with the ensemble sub-group mean (line) and range (shading) shown. MODIS data are the black dots.

In addition to runoff, the stability adjustment choice affects sublimation, with implications for the water (and glacial mass) balance. Switching on the adjustment approximately halves annual sublimation from 102 to 52 mm on average, with the mean relative reduction broadly consistent between input strategies. This means that sublimation accounts for around 9% and 5% of total catchment ablation with the stability adjustment switched off and on, respectively. However, further work is required to understand the substantial uncertainties and balance of

sensible/latent heat flux effects here, as well as the role of other processes, such as sublimation during wind transport.

Spatial Variation

The results above demonstrate that model configurations with diagnostic albedo, no drainage parameterisation and no stability adjustment, i.e. "0" options (Table 4.1), tend to induce earlier runoff at the catchment scale. Here we examine how these tendencies are manifest spatially and how the influence of different processes depends on both space and time, focusing on the crucial vertical dimension.

Figure 4.8 shows that S-shaped profiles of runoff differences develop and migrate upwards as the melt season progresses. These profiles form because early runoff at relatively high elevations using the "0" options leads to faster snow depletion relative to using the "1" options, leaving the latter with more mass to melt later in the season. This is consistent with the catchment responses described above.

These S-shaped profiles migrate upwards in sequence, with drainage followed by albedo and stability. Drainage option choice is particularly critical around the freezing isotherm for daily maximum temperatures, determining whether early melt is released or subject to storage through refreezing/melting cycles. In comparison, the lower elevation of peak albedo parameterisation effects is consistent with a more efficient LST-albedo feedback under slightly higher daily mean temperatures. The peak effects of stability adjustment choice are at lower elevations again, as they depend on the development of large near-surface temperature gradients. Notably, for both albedo and drainage, differences in runoff are present up to the highest elevations. Therefore, how liquid water processes and albedo are represented is critical for simulating high elevation perennial snowpacks and glacier accumulation, as well as the timing of runoff generated at lower elevations.

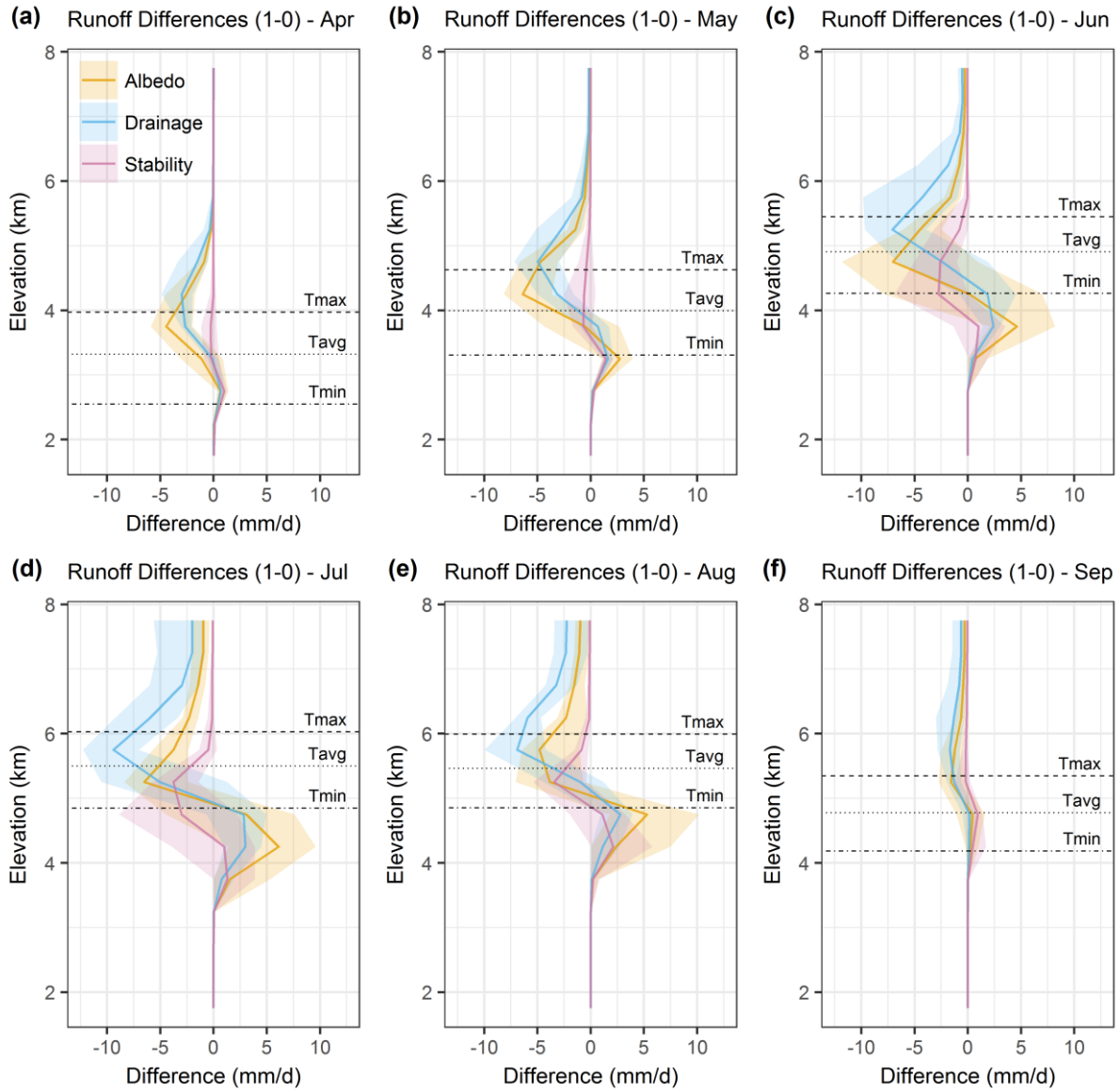


Figure 4.8 – Spatial (vertical) and temporal (monthly) differences in simulated runoff as a result of albedo, drainage and stability option choices. The differences are calculated in the same way as for Figure 4.4, but using 500 m elevation bands. Monthly mean freezing isotherm elevations for daily minimum, mean and maximum temperatures are also shown. The runoff rates are calculated based on elevation band areas, such that hypsometry is not taken into account. Results are based on FSM1 inputs.

4.4.4 Model Choice Interactions

In this section we examine how the tendencies of individual processes interact to shape the structure of the ensemble. This starts with the period mean case, which is followed by a consideration of inter-annual variability.

Period Mean

Figure 4.9a shows mean annual cumulative runoff curves for the ensemble grouped according to their combination of albedo and drainage option choices. This indicates that the fast runoff responses in the ensemble are associated with members employing the diagnostic albedo parameterisation in combination with instantaneous drainage. In contrast, the slowest runoff response occurs for ensemble members using prognostic albedo and the drainage parameterisation. In these bounding/end-member cases, the choices of albedo and drainage therefore act to exacerbate each other's individual tendencies to accelerate or slow runoff. However, the remaining two combinations of albedo and drainage options tend to result in similar cumulative runoff profiles, especially over the earlier part of the melt season. In this case, model choices can lead to a degree of compensation, i.e. combining the "fast" albedo option with the "slow" drainage option gives similar overall runoff results to combining "slow" albedo with "fast" drainage here.

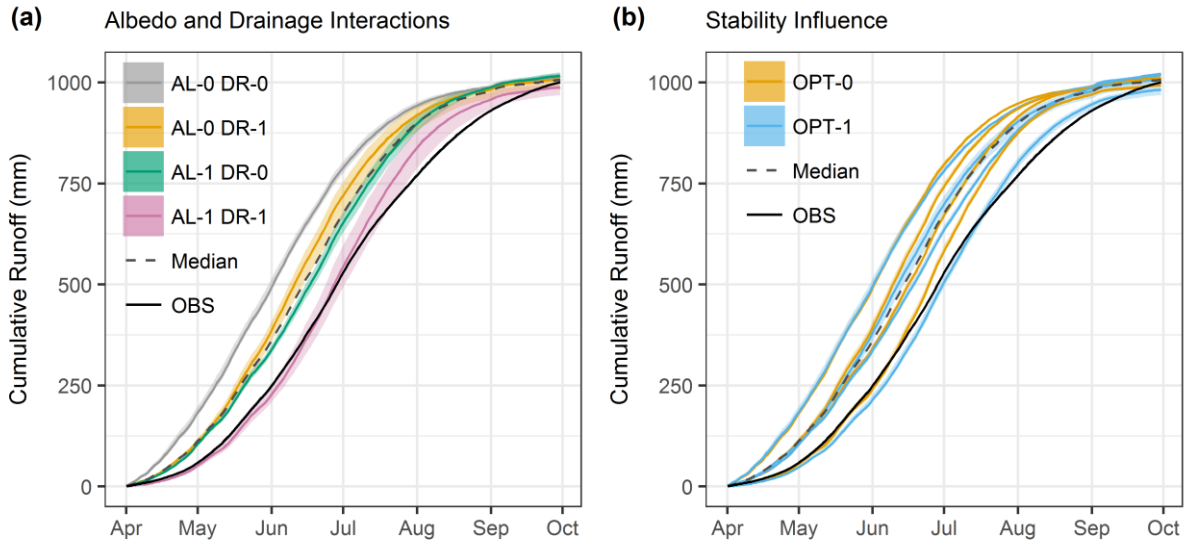


Figure 4.9 – Period mean cumulative runoff curves grouped by process options. Combinations of albedo (AL) and drainage (DR) options are shown in (a), with stability options additionally shown in (b). The latter panel is based on further differentiating the albedo/drainage combinations (shown in (a)) by stability option (i.e. 0 and 1). Both panels show mean annual cumulative runoff from observations (solid black line), as well as the ensemble mean (dashed line). Model results are based on FSM1 inputs.

Three of the albedo/drainage combinations in Figure 4.9a show increasing spread as the melt season progresses. Figure 4.9b confirms that this is due to the rising prominence of the

stability option choice later in the melt season, which becomes most prominent in the case of applying prognostic albedo together with the drainage parameterisation. Not applying the adjustment leads to more rapid runoff and earlier convergence with the other groupings. This fits with the explanation of the role of the stability adjustment above (Section 4.4.2), namely determining whether rapid melting can occur where snow is present under warm air temperatures. The observed mean cumulative runoff curve appears to be most consistent with the group applying prognostic albedo with the drainage parameterisation, but the correspondence with the stability option is more ambiguous. This is discussed further in Section 4.4.5.

Inter-Annual Variability

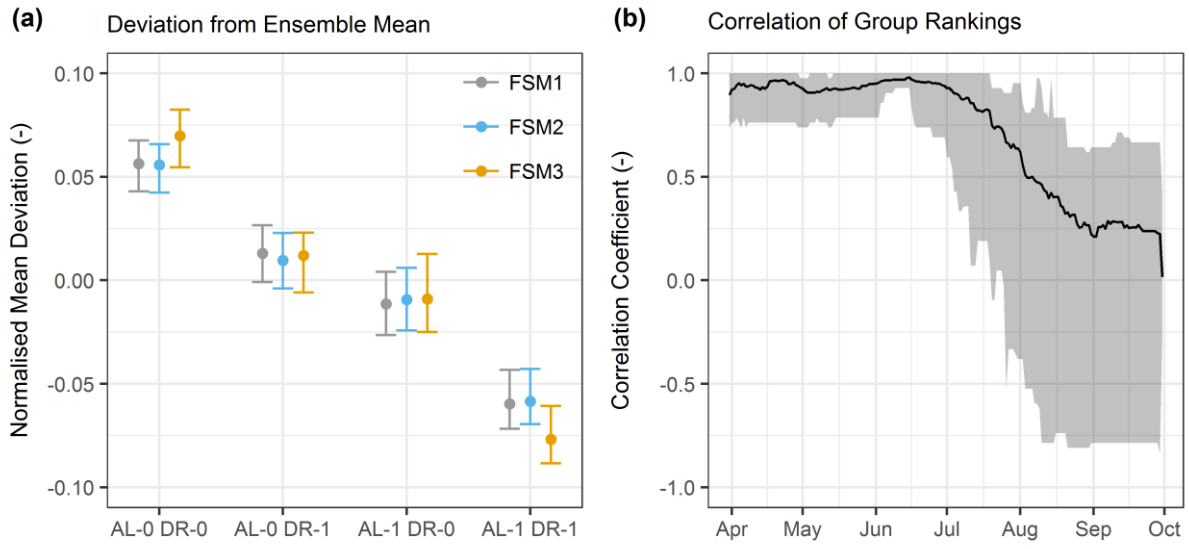


Figure 4.10 – Inter-annual variability in ensemble grouping structure. Mean deviations from the ensemble mean cumulative runoff curve (normalised to account for differences in annual runoff between years) for albedo (AL) and drainage (DR) option combinations are shown in (a) for each input strategy. Inter-annual variation in the rankings of groups within the ensemble compared with the period mean ensemble structure is shown in (b) as a time series of correlations (based on FSM1 inputs). The inter-annual mean (line) and range (shading) are plotted. The correlations are calculated for each day by aggregating the full ensemble on albedo, drainage and stability options, ranking the groups based on cumulative runoff and then correlating rankings with the period mean.

To see whether the pattern of groupings is consistent across years and input strategies, Figure 4.10a shows average deviations from the ensemble mean cumulative runoff curves for each albedo/drainage option grouping. Positive deviations indicate cumulative runoff generally above the ensemble mean curve and vice versa. From this we can see that the overall role of albedo and drainage option choices and their interactions is similar across input strategies and

between years. Diagnostic albedo applied without representation of liquid water results in the steepest cumulative runoff (i.e. positive deviation), while prognostic albedo with drainage results in the shallowest curve (i.e. negative deviation). The remaining two combinations are both close to the ensemble mean, reflecting the compensation effect noted above.

In Figure 4.10b we group the ensemble according to the possible combinations of albedo, drainage and stability options to look at how the group rankings correlate with the mean annual case over the course of the melt season. High correlations in the spring and early summer show that there is initially a consistent ranking across years. The correlation does deviate from unity, which is related to inter-annual variation in the two intermediate response albedo/drainage groups in Figure 4.9a and Figure 4.10a (Appendix B, Section B7). However, lower correlations from July onwards suggest more variable ensemble groupings later in the melt season. The significance of this for catchment runoff variations decreases as the end of the melt season approaches, as all groups approach similar cumulative runoff totals. Yet for some applications or critical zones these differences could be important, for example in perennial snow mass balance.

4.4.5 Group Performance

Section 4.4.4 thus show how different combinations of model options lead to exacerbation or compensation of their individual tendencies. This is driven primarily by albedo and drainage option choices at first, before the stability adjustment option also starts to play an increasingly significant role. The structure of the ensemble appears to be fundamentally similar in different years. We now consider how this relates to performance variation within the ensemble and between years.

Figure 4.11a shows RMSE for both cumulative runoff (for April to June, the major period of snow cover depletion) and SCA, with the ensemble grouped by albedo, drainage and stability options. From this we can see that the groups that perform better in terms of runoff also tend to show more consistency with MODIS SCA. This suggests that more accurate runoff response arises from better simulation of intra-catchment variation, rather than error compensation. Crucially, the best-performing groups all use the prognostic albedo option, which fits with the process-level evaluation against remote sensing and local data above (Section 4.4.2). The two groups that additionally include the drainage parameterisation show some of the best performance. These two findings are both in line with Essery et al. (2013) and Magnusson et al. (2015). However, the group omitting the drainage parameterisation but

applying the stability correction also shows low RMSE overall. This alludes to compensatory effects in this combination, with lower melt rates under the stability correction counteracting the rapid response induced by omitting the drainage parameterisation. More generally, applying the stability adjustment tends to reduce SCA errors, but improvements for runoff are small or mixed.

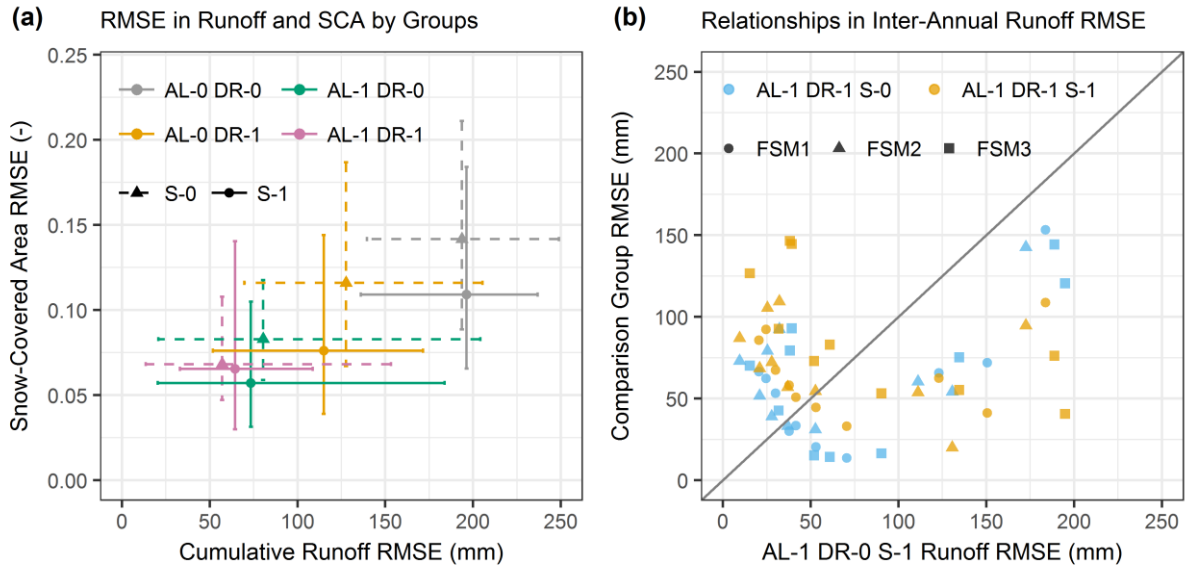


Figure 4.11 – Inter-annual variation in model performance by major groupings. Cumulative runoff RMSE is plotted against SCA RMSE in (a) for each of the albedo (AL) and drainage (DR) option combinations (i.e. averaging respective ensemble members), with differentiation by stability (S) option (shape and line type) also shown. SCA is based on an NDSI threshold of 0, consistent with the model classification of snow-covered cells as those containing any snow. Results are based on FSM1 inputs. In (b) the relationships between inter-annual variation in cumulative runoff RMSE for the three albedo/drainage/stability option combinations with the best overall performance are shown. RMSE for the two options incorporating the drainage parameterisation (AL-1 DR-1 S-0 and AL-1 DR-1 S-1) are differentiated by colour (blue and orange, respectively) and plotted against the next best configuration (AL-1 DR-0 S-1). Each of the input strategies are shown (different shapes).

Importantly, inter-annual variability in all groups is high (shown by the bars in Figure 4.11a), with notable overlap between groups. To examine this further, we investigate the relationship between performances of different groups in Figure 4.11b. Focusing on runoff RMSE for the three best-performing groups overall, we can see that trade-offs resembling a Pareto front develop. This means that, for a number of the years simulated, performance in one group cannot increase without a corresponding reduction in performance in another group. Specifically, for the groups in Figure 4.11b, years with good performance with the drainage option off (and the stability adjustment applied) are associated with performance reductions in

configurations where the drainage option is switched on. This indicates that a model structural constraint on performance exists. As shown above, the tendencies from different parameterisation options and their interactions are essentially consistent between years. However, this likely limits the extent to which a single configuration can prove optimal in all years or for all variables of interest.

This latter point builds on the findings of non-uniqueness in previous model inter-comparison studies (see Section 4.1) by showing that performance trade-offs are systematic for the best groups of model configurations in realistic applications. In Appendix B (Section B8) we provide more such inter-group comparisons, which show the range of relationships that exist when considering groups with a broader range of performance.

There are of course limitations with this performance evaluation, which stem from modelling simplifications (Section 4.3.1) and input/evaluation data uncertainty (Section 4.3.2). However, the consistency between SCA and runoff performance metrics noted above does add confidence. Indeed, while absolute RMSEs might depend on the approaches taken here, their relative rankings at least are likely to be robust and consistent with physical reasoning. Moreover, the inter-annual and inter-group relationships in performance are shown to be consistent across input strategies.

4.4.6 Climate Sensitivity

To contextualise the significance of these findings, we briefly consider the simple climate sensitivity test described in Section 4.3.4. Figure 4.12 quantifies how monthly runoff in the two end-member albedo/drainage combinations (referred to here as the "fast" and "slow" combinations) responds differently to temperature perturbations. The "fast" combination shows a heightened sensitivity in April, which drops below that of the "slow" combination in May. By June, the "fast" combination actually shows a generally negative relationship between runoff and temperature perturbations, reflecting the fact that increased melting earlier in the season leaves less snow available during the peak energy inputs of summer. It takes longer for the "slow" combination to reach this situation, but by July both combinations show similar responses. Moreover, while runoff sensitivity is approximately linear on average, varying degrees of nonlinearity are present in individual years, particularly from May onwards. This is influenced by inter-annual variability in winter and spring snowfall through its control on how early runoff may become strongly mass-constrained.

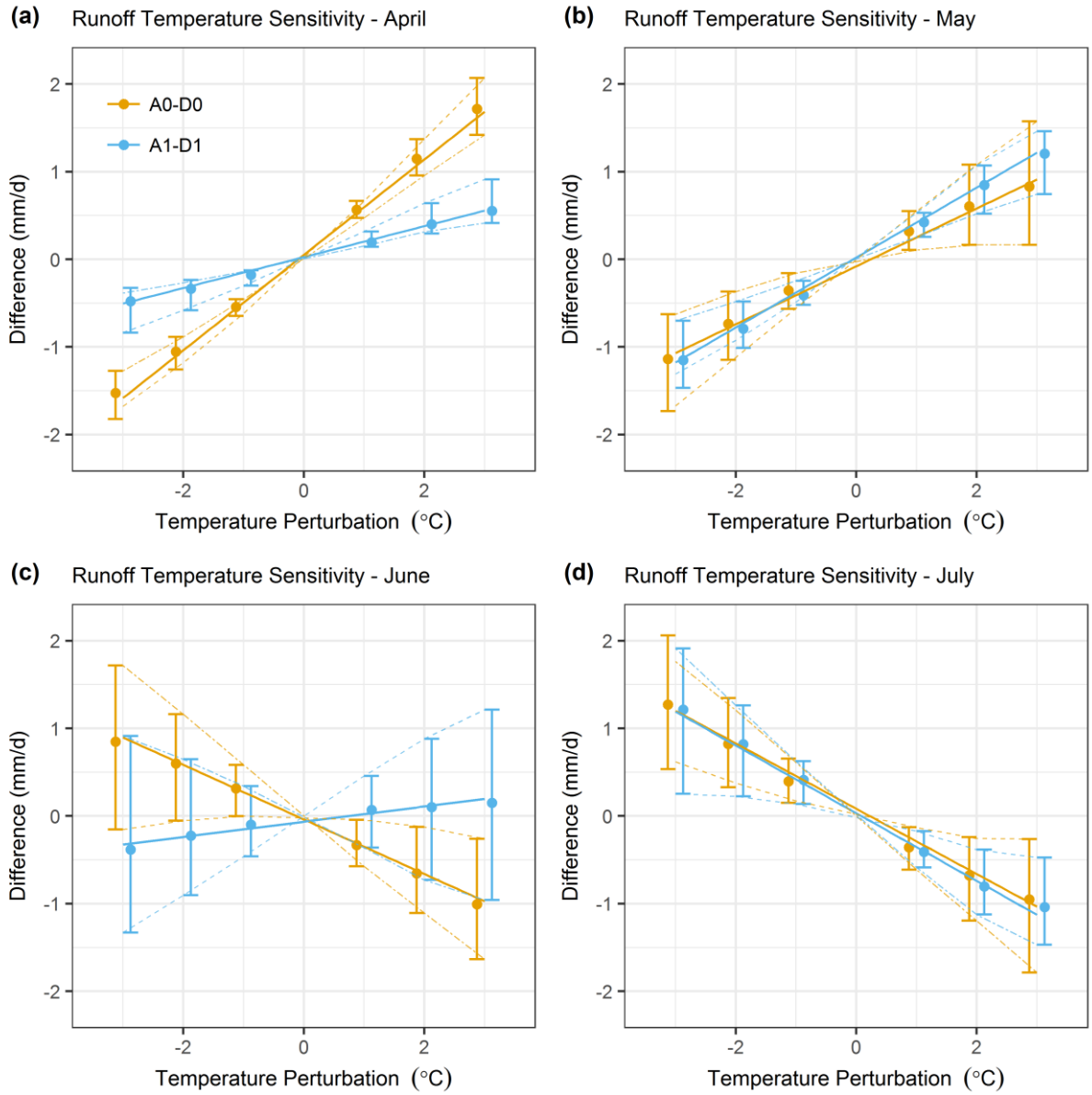


Figure 4.12 – Differences in monthly runoff arising from summer (AMJJA) temperature perturbations for the two end-member ensemble groupings (based on albedo (A) and drainage (D). The "fast" group (A0-D0) is shown in orange and the "slow" group (A1-D1) is shown in blue. Dots denote the mean response across years, with simple linear regression lines (solid) shown. For clarity, the dots are slightly offset from the specified temperature perturbations on the bottom axis. Individual years corresponding with the minimum and maximum responses are also plotted using dashed lines. Differences are relative to the baseline (FSM1) inputs, i.e. without any perturbation.

While more detailed analyses could be undertaken in future studies, these results have several important implications. For example, the "fast" albedo/drainage combination could lead to more rapid and pronounced shifts in runoff timing in hydrological projections, with larger changes at higher elevations also likely. This combination might also be expected to amplify the snow-albedo feedback that contributes substantially to spread in climate model

simulations (e.g. Qu and Hall 2014), which would strongly affect projections of elevation-dependent warming (Pepin et al. 2015; Palazzi et al. 2017). Further investigation is needed to see how the magnitude of this tendency might change following parameter calibration, although the "fast" combination does appear intrinsically less realistic (Sections 4.4.3-4.4.4). Finally, we note that simpler snow models (e.g. temperature index) might respond differently to climate perturbations depending on how they represent (implicitly or explicitly) or lump processes related to albedo, storage/drainage and turbulent fluxes. This also requires further investigation.

4.5 Conclusions

This study finds that using HAR-based inputs with a spatially distributed version of the FSM leads to generally good model performance at the ensemble level in the Himalaya. Based on validation of both the HAR and model response against local observations and multiple remote sensing products (Chapter 3), this finding confirms the substantial potential for high resolution regional climate modelling to supplement other data sources in the Himalaya and support snow model applications. However, we also show that ensemble spread is significant and constituted by groupings with different levels of model performance. These different responses of FSM ensemble members are strongly controlled by how processes related to albedo and liquid water storage/drainage are represented, with a secondary influence from turbulent fluxes (i.e. stability adjustment). Consistent tendencies to accelerate or delay runoff arise depending on the option selected, but the relative importance of different processes depends on both space and time. In combination, these tendencies may have exacerbating or compensating effects, but these lead to a similar overall ensemble structure between years, especially in the first half of the melt season.

From evaluation using local data and remote sensing products, we find that configurations with better performance for runoff also more accurately reproduce SCA. In addition, the results suggest that the prognostic albedo parameterisation should be consistently preferred, in agreement with other studies (Essery et al. 2013; Magnusson et al. 2015). The liquid water storage/drainage option is also applied in most of the well-performing configurations, which is physically realistic, but some compensatory effects arise depending on the stability adjustment option selected. Based on comparisons with MODIS LST, there is some suggestion that turbulent fluxes are overly suppressed by applying the stability adjustment. This is plausible based on previous studies (see e.g. Andreadis et al. 2009), but whether one

stability adjustment choice should be preferred is difficult to determine fully. However, the approach of using multiple datasets and process-based modelling allows us to identify such uncertainties and formulate possible responses. As discussed in Section 4.4.2, it would be useful in further work to test additional options for stability adjustment and possible missing processes, including terrain enhancement of incoming longwave radiation and sensible/latent heat advection. Regional climate modelling with higher resolutions than the HAR is desirable here, as it may well improve relevant inputs further, although more in-situ climate and snow measurements are needed to really quantify this.

The findings above suggest that there are structural limits on model performance. As the process options essentially behave and interact similarly between years, it is difficult for a single model to capture all of the inter-annual variability. Trade-offs in performance may thus ensue. Furthermore, we show that different configurations can exhibit substantially different climate sensitivities, although more detailed analysis is required in further work. Together, these points suggest that the ensemble modelling approach should be retained in applications, including climate change impact assessment. In some cases at least, it could be reasonable to use a subset of the ensemble. This could be based on removing some members where processes have less influence on variables of interest, but particularly where good reasons exist to suggest that particular parameterisations are less appropriate (e.g. diagnostic albedo). Where uncertainties are larger, for example for stability adjustment, it makes sense to retain multiple options. This could help to better characterise uncertainty, which is urgently required to support water resources management in the Himalaya in a changing climate.

Chapter 5

Comparison of Cryospheric-Hydrological Modelling Approaches

5.1 Introduction

The broad field of hydrological modelling has long been characterised by two competing modelling philosophies (Hrachowitz and Clark 2017). On the one hand there are detailed, typically high resolution models based on the small-scale equations best describing catchment processes. On the other hand there exist myriad conceptual, often spatially lumped and simpler formulations. Models in the latter class use functions to relate catchment inputs and outputs, without explicit regard to the physics underpinning hydrological processes. Along the continuum between these model types, substantial variation in complexity and resolution is found. Opinions vary on appropriate process descriptions at different scales, as well as suitable complexity in the face of partial data (e.g. Beven 1989; Grayson 1992; Beven 2001, 2002, 2006; Wood et al. 2011; Beven and Cloke 2012; Montanari and Koutsoyiannis 2012; Fatichi et al. 2016; Savenije and Hrachowitz 2017). Yet, in seeking strategies to improve representations of hydrological processes and their scaling behaviour, the complementarity of contrasting models and the importance of systematic inter-comparison have recently been re-emphasised (Hrachowitz and Clark 2017; Clark et al. 2017; Peters-Lidard et al. 2017).

For the mountainous Upper Indus Basin (UIB), where the majority of river flow is derived from snow and ice melt, questions regarding the applicability of different hydrological models centre on representations of cryospheric processes. As exemplified by commonly applied ablation algorithms, severe data paucity has often led to simpler approaches being preferred across the Himalayan arc. These are typically versions of temperature index (TI) (e.g. Bocchiola et al. 2011; Rees and Collins 2006; Immerzeel et al. 2009, 2012a, 2013; Soncini et al. 2014) or enhanced temperature index (ETI) approaches (e.g. Ragettli et al. 2013, 2015; Immerzeel et al. 2014; Minora et al. 2015). The former scales melt with air temperature following exceedance of a critical temperature threshold, while the latter additionally incorporates a term for net shortwave radiation at the surface. Many of the aforementioned studies consider these approaches to perform satisfactorily with respect to observed river flows, remote sensing of snow cover dynamics and, increasingly, regional glacier mass balance estimates from geodetic or other methods.

As essentially conceptual formulations abstracted from the surface energy balance, one of the outstanding questions with respect to these melt modelling approaches is their transferability in space and time to conditions outside those experienced during calibration. This is highly relevant for hydrological projections in a changing climate. While it has long been established that TI parameters vary notably in space and time (e.g. Lang and Braun 1990; Hock 2003), Gabbi et al. (2014) found there to be substantial implications of temporal parameter variation for relatively long-term simulations of glacier mass balance. From their results, an ETI model was suggested to be more robust, although other studies question whether this finding is general. MacDougall et al. (2011) noted that optimising an ETI model (Pellicciotti et al. 2005) in their Canadian study area actually reduced it to a simple TI model. In the French Alps, Vincent and Six (2013) and Réveillet et al. (2017) argued that accounting for shortwave radiation variability is certainly important to help capture spatial variation in melt, but temperature variation primarily controls inter-annual variability and glacier-scale response.

For the latter two studies, the findings may be partly guided by their emphasis on clear-sky (potential) rather than cloud-modulated shortwave radiation. The reasons behind MacDougall et al.'s (2011) findings were less clear to the authors, but possibly reflect climate-dependent performance of the ETI model. Indeed, Carenzo et al. (2009) found that ETI performance deteriorated under overcast conditions in the Swiss Alps, while the parameters associated with the temperature and net shortwave radiation components of the model varied spatially with wind regime and the relative importance of turbulent fluxes. Matthews et al. (2015) also showed that ETI parameters can be considered temporally variable, such that making model parameters a function of weather type improved performance for both TI and ETI models.

An alternative to attempting to account for spatial and temporal parameter variation in TI or ETI models is to apply distributed surface energy balance (EB) modelling. This approach provides a closer approximation to the physics driving ablation, which has been shown to help delineate contrasting responses to climate changes in different climate-snow regimes (e.g. López-Moreno et al. 2017). Yet, distributed EB modelling in the Himalaya comes with its own challenges, not least in the form of supplying appropriate climate inputs. Various studies have demonstrated the performance drop due to forcing energy balance models with non-local or estimated climate data, particularly on glaciers (e.g. MacDougall and Flowers 2010; Gabbi et al. 2014; Réveillet et al. 2018), while turbulent fluxes, small-scale variability and unresolved processes contribute notable uncertainties. However, there is now the possibility that high resolution coupled land-atmosphere model outputs can be used to force such models

with more confidence, even in data-sparse areas. This approach is still relatively unexplored in the Himalayan region, although Chapters 3/4 demonstrate its clear potential, while other studies highlight the increasing skill of coupled land-atmosphere modelling in the region (Collier et al. 2013, 2015). As such, it seems reasonable to ask how such an approach would compare with more commonly used conceptual formulations like the TI and ETI models.

While comparisons of TI, ETI and EB models have been undertaken previously, these have been typically for non-Himalayan climate, snow and glacier regimes (e.g. MacDougall et al. 2011; Gabbi et al. 2014). This is partly due to available evaluation data for ablation processes being limited, indirect and somewhat noisy, with persistent uncertainties additionally remaining even in new sources of climate input data, such as the High Asia Refined Analysis (HAR) (Maussion et al. 2014). However, we contend that such a model comparison in the UIB may usefully consider differences in model response and climate sensitivity, as well as performance in relation to observations as far as is possible. This is because, for practical hydrological projections in a changing climate, the differences in sensitivity of alternative input/model combinations to climate variability and perturbations are of substantial interest. Even if a model appears to give good performance in typically short calibration and/or validation periods, this is no guarantee that it approximates climate sensitivity well enough for longer projections (e.g. Gabbi et al. 2014). This is particularly the case for glacier mass balance trajectories, where the cumulative effects of differing climate sensitivities could lead to substantial divergence.

Within this context, the chapter aims to evaluate different cryospheric and hydrological modelling approaches in the UIB, focusing primarily on snow and glacier ablation. This is achieved in two parts. Firstly, the TOPKAPI-ETH model is set up and analysed for snow- and glacier-dominated sub-basins of the UIB. The purpose of this is to provide a performance benchmark for the ETI-centred approach, while also giving further validation of the climate inputs used. TOPKAPI-ETH has been successfully applied to partly glacierised catchments in the European Alps (Finger et al. 2011), the Karakoram (Pellicciotti et al. 2012; Ragettli et al. 2013), the Nepalese Himalaya (Ragettli et al. 2015) and the Andes (Ragettli and Pellicciotti 2012; Ragettli et al. 2014, 2016). Secondly, based on modifications and extensions to the Factorial Snowpack Model (FSM) (Essery 2015) to provide a framework for systematic ablation model inter-comparison, the chapter analyses TI, ETI and EB approaches to both snow and glacier melt modelling. The FSM extensions introduced here are intended ultimately to augment the existing open source framework for directly comparing different

model strategies, minimising the confounding influences of varying model component implementations between programs. The extended version of FSM coded in this work is denoted as FSM+. The analysis explores how spatial and temporal variability in climate and cryospheric processes leads to differences between the models.

5.2 Data and Methods

This section begins by describing the input and evaluation data used in the models of the Astore and Hunza sub-basins in this chapter. The Astore and Hunza are representative of snow-dominated and heavily glaciated hydrological regimes, respectively. Thereafter, the TOPKAPI-ETH model is introduced, along with the benchmark configuration analysed here. Finally, the section explains additional developments to the spatially distributed version of FSM introduced in Chapter 4 (referred to here as FSM+).

5.2.1 Climate Inputs

The climate inputs used in this chapter are based on the data sources and processing methods reported in Chapter 4 for the spatially distributed snow process simulations of the Astore sub-basin. As such, the HAR (Maussion et al. 2014) is used as the basis for spatiotemporally varying input fields of rainfall, snowfall, air temperature, relative humidity, wind speed, surface air pressure, and incoming shortwave and longwave radiation. Following preliminary testing of different strategies, a minimal bias correction approach was adopted. Specifically, we corrected near-surface air temperatures for biases in the mean, based on local observations. The reasons for this approach are outlined in full in Chapter 4. For the Hunza sub-basin, precipitation was also adjusted by a constant factor (0.8), which was estimated based on the mean annual water balance considerations for the HAR in Chapter 3. While FSM+ utilises all of the aforementioned climate variables to solve the surface energy balance, TOPKAPI-ETH only requires precipitation, temperature and cloud transmissivity. The latter variable is used to factor clear-sky incoming shortwave radiation for the TOPKAPI-ETH's ETI component.

The same methods were also used for disaggregation of the HAR to the higher resolution needed in this modelling. These methods draw on the approaches within the MicroMet (Liston and Elder 2006b) meteorological pre-processor of SnowModel (Liston and Elder 2006a), as well as the methods used by Duethmann et al. (2013). The method uses linear regression to relate each variable to elevation, the primary control on climate variation in the

UIB, with prior log-transformation for precipitation. The values at each HAR cell are then interpolated to a reference level using the gradient determined in regression. This gives a spatial (horizontal) anomaly field, which is interpolated to high resolution, and then the elevation signal is reintroduced using the regression gradient. This helps to capture the most critical sub-regional variation at scales resolved in the HAR, but not more local, topographically modulated patterns. Where regression is not statistically significant at the 95% interval, simple interpolation of the HAR input field is used. Moreover, the probable decrease in precipitation at very high elevations, above 5500 mASL, is handled simply by supposing that an elevation (z) above this threshold can be treated as $5500 - (z - 5500)$ for the purposes of precipitation disaggregation only. The only variable handled differently is incoming shortwave radiation, which is modelled following Corripio (2003), modulated by cloud transmissivity estimated from the HAR (see Chapter 4).

Based on the results in Chapter 4, this approach appears to be adequate at least for the purposes of obtaining a plausible enough set of climate forcings for evaluating different cryospheric and hydrological modelling approaches. Nevertheless, there are of course a number of limitations present. These include challenges in quantitative bias correction from a small number of station locations measuring only selected variables, as well as accounting for topographic influences other than elevation. However, to some extent this represents a fairly typical application in the Himalayan region at present. While this makes it harder to make totally general statements regarding performance of different model types, it does help to provide some insight into how different models compare in a realistic case.

To provide longer time series for model input, the HAR-based precipitation and temperature series were extended based on anomaly series derived from long-term observations. This helps to combine the information provided by the HAR's good spatial resolution with the better temporal coverage of observed series. Monthly anomaly time series for the period 1961-2010 were calculated from in-situ observations, which were then used to estimate absolute series for the HAR grid cells in the NWUIB for the same period. Testing showed that the correlation of anomalies between HAR grid cells in the domain was very high on a monthly and seasonal basis, such that it was not considered necessary to account for any further sub-domain variability. The estimated HAR monthly fields were then disaggregated in time using the observed daily time series as a guide to daily precipitation fractions and temperature quantiles. Sub-daily disaggregation was not undertaken, as the extended series are only applied with the TOPKAPI-ETH model, which was run using daily inputs only. The

efficacy of this method is demonstrated by the results in Section 5.3.1, although of course spatial and temporal variability is simplified to a degree. Simulations with the augmented FSM program are restricted to the HAR record period.

5.2.2 Static Inputs

The static inputs used are consistent between the TOPKAPI-ETH and FSM+ models. All static inputs are on the 500 m resolution model grid in a UTM43N projection. The digital elevation model (DEM) is based on the Shuttle Radar Topography Mission (SRTM) V4.1 data product (Jarvis et al. 2008), while initial glacier area distributions are taken from the Randolph Glacier Inventory (RGI) V5.0 (Arendt et al. 2015). Maps of initial glacier thickness were estimated based on the equilibrium shear stress approach (Cuffey and Paterson 2010b), but climate inputs were cycled to reach an approximate dynamic equilibrium in the relevant model runs. This is clearly an over-simplification, although glacier mass balance, length and area have been reasonably stable in the Karakoram for several decades (Bolch et al. 2017; Zhou et al. 2017). Debris cover on glaciers is based on the discrete classification produced from Landsat imagery by Khan et al. (2015). The admittedly coarse distribution of soil types is taken from the Harmonised World Soil Database V1.2 (Fischer et al. 2008), while land cover is taken from the 0.05° ESA GlobCover 2009 product (Arino et al. 2012).

5.2.3 Evaluation Data

Available evaluation data include daily river flows recorded by the Water and Power Development Authority (WAPDA) for the Doyian (Astore) and Dainyor (Hunza) gauging stations. No local snow measurements are available, but we compare model outputs with snow-covered area SCA derived from the MODIS MOD10A1 product (Collection 6) (Hall and Riggs 2016), as a further spatial constraint on internal catchment processes. As cells are classified as snow-covered in TOPKAPI-ETH and FSM+ if any snow is present, MODIS SCA was determined using a normalised differential snow index (NDSI) threshold of 0.0. We also explore the potential for the MODIS land surface temperature (LST) (MOD11A1) product (Collection 6) to support evaluation of processes affecting the surface energy balance for glaciers. The equivalent analysis for snow cover was presented in Chapter 4. The LST products have been widely evaluated (Liu et al. 2009; e.g. Wan et al. 2004) and, as discussed in Chapter 4, additional evaluation using local observations at the Concordia site is given in Appendix C. This confirms that the MODIS product generally performs well, with low enough bias to indicate gross model errors.

5.2.4 TOPKAPI-ETH

Background

The TOPKAPI (TOPographic Kinematic APproximation and Integration) hydrological model was originally developed to improve upon the semi-distributed conceptual models ARNO (Todini 1996) and TOPMODEL (Beven and Kirkby 1979). ARNO is structured around simple functions governing soil moisture storage. TOPMODEL is based on a topographic index relating upstream drainage area to the local slope in a DEM, as well as a negative exponential relationship between water table depth and soil transmissivity. In both cases, some of the model parameters are not physically meaningful, at least at the scales required in catchment modelling (Todini 1995; Ciarapica and Todini 2002). This potentially inhibits their application to ungauged catchments, and under changing land use or climatic conditions. In addressing these issues, TOPKAPI adopts a more process-based, spatially distributed approach, accounting for variation in theoretically measurable catchment properties.

At the core of TOPKAPI is its application of kinematic wave theory to soil, overland and channel flow routing. The structure of the resulting nonlinear reservoir equations is similar for these three components, all assuming that the role of topographic slope is predominant (Todini 1995; Ciarapica and Todini 2002; Liu and Todini 2002). This strategy exemplifies the intermediate position of TOPKAPI on the spectrum of hydrological model complexity. It exhibits reasonable process fidelity, especially in certain contexts, but without some of the additional accuracy, data requirements and computational costs of more advanced physically based models, such as SHETRAN (Abbott et al. 1986; Ewen et al. 2000). In theory, each of the required parameters in TOPKAPI can be derived from soil and land use maps and databases, although some calibration may of course still be required, partly due to the persistent challenges in mapping catchment heterogeneity, as well as inherent uncertainty and noise in hydrological data. With extension to include deeper soil layers, infiltration, percolation and groundwater flow, the TOPKAPI flow routing approach was shown to perform well in flood forecasting (Liu et al. 2005) and various geographical contexts (e.g. Ciarapica and Todini 2002; Liu and Todini 2002; Sinclair and Pegram 2010).

Developments at ETH-Zurich

The core features of TOPKAPI have been retained in the TOPKAPI-ETH variant, which represents an adaptation of the original model for improved representation of mountain region

processes (Finger et al. 2011). While linear reservoirs handle the routing of glacial meltwater in TOPKAPI-ETH, the fundamental kinematic wave approaches to soil, overland and channel runoff production and flow routing remain. Rather, the major changes to the original model are in representation of snow and glacier processes, including incorporation of an enhanced temperature-index (ETI) method to calculate melt (Pellicciotti et al. 2005). The ETI method is defined as

$$M = \begin{cases} TF \cdot T + SRF \cdot (1-\alpha) \cdot G, & T > T_T \\ 0, & T \leq T_T \end{cases} \quad (1)$$

where M is melt (mm w.e. hr^{-1}), TF is the temperature factor (mm hr^{-1} °C^{-1}), T is the temperature (°C), SRF is a shortwave radiation factor (m^2 mm W^{-1} hr^{-1}), α is albedo, G is incoming shortwave radiation (W m^{-2}), and T_T is the threshold temperature for melt onset (°C). Incoming shortwave radiation under clear sky conditions in TOPKAPI-ETH is modelled following a vectorial algebra approach (Corripio 2003), which is then modified by a time-varying and user-specified cloud transmissivity. Snow albedo is parameterised using a logarithmic function based on the cumulative sum of daily maximum positive air temperatures since the last snowfall (Brock et al. 2000). Ice albedo is left constant, while alternative temperature and shortwave radiation factors may be specified where debris cover is present on glaciers (referred to as ETId hereafter) (e.g. Ragettli et al. 2013).

An adapted version of the ETI formulation was also recently developed to account for debris thickness, rather than just debris cover, albeit at the expense of introducing a number of additional calibration parameters. Specifically, the ETI model accounting for debris thickness (DETI model hereafter) introduced by Ragettli et al. (2015) and Carenzo et al. (2016) takes the form:

$$M = \begin{cases} TF_d \cdot T(t-lag) + SRF_d \cdot (1-\alpha) \cdot G(t-lag), & T > T_T \\ 0, & T \leq T_T \end{cases} \quad (2)$$

where TF_d is the temperature factor (mm hr^{-1} °C^{-1}), SRF_d is a shortwave radiation factor (m^2 mm W^{-1} h^{-1}), t is the current time step (hr) and lag is a delaying parameter (hr). This formulation thus parameterises transmission of temperature and shortwave radiation signals through debris cover by lagging the climate input variables. Moreover, the TF_d and SRF_d are also made dependent on debris thickness in the DETI formulation, based on the relationships obtained by Carenzo et al. (2016) from field measurements and an energy balance model

accounting for heat conduction through debris (referred to as the DEB model) (Reid and Brock 2010). This introduces six auxillary parameters that are required to calculate the TFd, SRFd and lag parameters in equation 2, which clearly increases the dimensionality of the parameter estimation problem relative to the ETId method. It is not yet established how such an approach might fare in the Karakoram when using direct parameter transfer, as might be reasonable in a context with limited appropriate data for parameter re-estimation.

TOPKAPI-ETH also accounts for some key processes of snow and ice mass redistribution., including the aggregate effects of avalanches through the SnowSlide algorithm (Bernhardt and Schulz 2010). This simple approach defines a snow holding depth as an exponential function of the slope between neighbouring model cells that contain snow. Snow in excess of this threshold depth is transferred to the downstream cell. The model also approximates glacier mass redistribution using the Δh-parameterisation (Huss et al. 2010). The parameterisation proceeds by rescaling annual net mass balance by expressing normalised ice thickness change as a function of normalised elevation range. This method is essentially empirical; it requires parameters estimated from historical observations and is applicable to groups of glaciers. Nevertheless, Huss et al. demonstrate that it can perform well in comparison to 3D numerical models solving the nonlinear Stokes equations. As such, it represents one practical solution in areas where the data required for complex ice flow modelling is unavailable, although uncertainties over the transferability of empirical parameters to different regions may compromise its simplicity to some degree.

Potential evapotranspiration in TOPKAPI-ETH is estimated using the Priestley-Taylor method, with crop factors varying seasonally and with land use to control demands on storage in the upper soil layer and at the surface. Infiltration is based on an explicit solution of the Green-Ampt equation (Salvucci and Entekhabi 1994), while the percolation rate increases with soil water content following a power law (Clapp and Hornberger 1978).

Setup

The TOPKAPI-ETH baseline is based largely on the configuration described in Ragettli et al. (2013). Critically, the parameters related to snow and ice melt, snow redistribution by gravity, glacier mass redistribution and glacial meltwater routing are all taken from Ragettli et al.'s setup. Parameters controlling soil, overland and channel routing were fixed based on cross-referencing soil/land use maps and inferences on channel characteristics with standard tables of hydraulic properties (Maidment 1993). The climate and static inputs follow those outlined

in Sections 5.2.1 and 5.2.2, with the use of the HAR as the basis for climate inputs being the most important difference compared with Ragettli et al.

5.2.5 FSM+

Snowpack Processes

Snowpack processes in FSM+ are represented using the FSM multi-physics approach, which was adapted simply for spatially distributed simulations, as described in Chapter 4. To recap, the FSM multi-physics ensemble employed here is an intermediate complexity, systematic framework for testing alternative representations of key snowpack processes and how they interact with each other within a coupled mass and energy balance scheme (Essery 2015). With two parameterisation options (0/1) for five different snowpack processes, the FSM ensemble includes 32 possible model configurations. The parameterisation options neatly synthesise common approaches in a range of widely applied snow models. Analyses of FSM to date have shown that it gives ensemble performance and spread comparable to larger multi-model ensembles (Essery 2015). Its value for testing new process representations has also been demonstrated (Moeser et al. 2016).

Based on the evaluation of FSM for both the Concordia site and the snow-dominated Astore sub-basin in Chapter 4, the FSM ensemble is substantially restricted for the results presented in this chapter. Unless otherwise indicated, the members retained here use prognostic albedo, the liquid water retention and drainage parameterisation, prognostic density and variable snow thermal conductivity. Default parameter values from Essery (2015) are retained. The results in Chapter 4 suggest that the density and thermal conductivity representations may be of relatively minor importance where a study's emphasis is primarily hydrological in nature. As well as being physically more realistic, the albedo and liquid water parameterisation choices generally lead to better model correspondence with observations. However, there are several complicating issues discussed at length in Chapter 4, including temporal variation in the ostensible optimality of any given model configuration.

Another relevant finding from Chapter 4 is that selecting the most appropriate atmospheric stability adjustment option for the turbulent heat flux calculations is somewhat complicated. Comparisons with observed river flows, as well as MODIS SCA and LST, suggested that switching the stability adjustment on could lead to excessive dampening of turbulent fluxes. While this could partly represent the influence of unresolved processes, such as local (sub-

grid) sensible heat advection due to patchy snow cover, the issue of appropriately accounting for stability conditions has been found in other EB applications for snow (e.g. Jordan 1991; Tarboton et al. 1995; Andreadis et al. 2009) and glacier modelling (e.g. Conway and Cullen 2013). In light of this, Collier et al. (2015) elected to restrict the maximum damping of turbulent fluxes in stable conditions to 30% of the correction factor calculated by the bulk Richardson number approach. For this chapter, we decided instead to retain both the stability adjustment options in FSM, i.e. no adjustment and the unaltered bulk Richardson number method, in order to examine the differences between these two endpoints.

Glacier Accumulation and Ablation

A relatively simple approach to simulate clean ice ablation on glaciers was developed to fit in with the modular structure of FSM. For ablation calculations, the model considers the upper 10m of ice where glaciers are present. This depth was selected to match the approximate limit of penetration of the signal from the annual surface temperature cycle in glacial ice (Cuffey and Paterson 2010a), a commonly employed strategy in process-based models of glacier ablation (e.g. Mölg et al. 2008, 2009; Collier et al. 2015). Heat conduction through the glacier ice was modelled following the implicit scheme formulated in FSM, with vertical discretisation based on progressively increasing layer thicknesses (from 0.1 to 1 m) at successively greater depths. A fixed temperature was prescribed at the base of the simulated ice column, again following Mölg et al. (2008, 2009) and Collier et al. (2013, 2015). The temperature was set to the mean annual air temperature estimated from observations (and lapse rates derived from observations), up to a maximum of 273.15 K for any low-lying, typically heavily debris-covered glacier tongues. As a one-dimensional approach, with each grid cell independent of all others, this is clearly a simplification of the complicated, spatially and temporally varying temperature distributions likely to be found in UIB glaciers, given their vast elevation range and patterns of motion. Nevertheless, the aforementioned studies suggest that it provides a reasonable first approximation.

Glacier ice ablation is calculated using the surface energy balance approach, as in FSM, albeit with parameters appropriate for glacier ice. This includes albedo and aerodynamic roughness length, which were initially set at 0.35 (Cuffey and Paterson 2010a) and 0.001 m (Collier et al. 2015), respectively, although we note that literature estimates for the latter vary over one or more orders of magnitude (Brock et al. 2006). Glacier ablation does not occur in the model until any overlying snowpack has been completely removed. Similar to the case of snowpack

directly overlying a presumed soil substrate, heat conduction between the base of the snowpack and the upper glacier ice is calculated explicitly, as per FSM. While only the upper 10m of a glacier are considered for ablation calculations, full glacier thicknesses are tracked for assessing glacier mass balance. As described in Section 5.2.2, initial glacier thicknesses were estimated from the equilibrium shear stress approach (Cuffey and Paterson 2010b), followed by cycling climate inputs to attain an approximate dynamic equilibrium.

Uncertainties regarding firn processes and snow-ice transformations in the UIB are very high, with the additional complication that many glaciers are nourished by avalanching, such that "typical" accumulation zones are from ubiquitous (Hewitt 2014). This issue is set against a backdrop of ongoing challenges in modelling the moisture-dependent firn densification processes in mountain hydrological models. Forsythe (2012) used the parameterisation implemented by Saloranta (2012), noting that even then an empirical correction was required to keep the snow-ice transformation from occurring on unrealistic timescales. This is potentially an area requiring further development, but for now we adopt the simple approach to glacier ice accumulation in Soncini et al. (2014), where 10% of remaining snow mass in glaciated cells is transformed directly to glacier ice at the end of each melt season.

Debris-Covered Glacier Ablation

To account for the effect of extensive debris cover over parts of major glaciers in the Karakoram, a model of heat conduction through debris and sub-debris melt was also implemented in FSM+. This model is similar to the DEB formulation put forward by Reid and Brock (2010), as well as one of the model structures examined in Collier et al. (2014). Collier et al. compared the effects of including moisture and phase changes within debris on melt rates of underlying ice relative to a model omitting these influences (i.e. assuming "dry" debris). The model experiments indicated there to be some influence on the exact partitioning of surface energy balance terms, albeit with seemingly little net effect on simulated surface temperatures, with the omitted negative latent heat flux in the "dry" case being offset by a heightened (negative) sensible heat flux. Including a moisture parameterisation led to some difference in sub-debris melt when viewed cumulatively over the simulated ablation period, but the magnitude of these differences was reasonably small (3.1-7.5%), especially in the context of the errors and approximations inherent in UIB modelling.

As such, we opted to use a formulation similar to the "dry" debris case for simplicity. That is not to say that the moisture effects are insignificant, especially for longer term glacier

evolution. Indeed, Collier et al. (2015) usefully take forward the model accounting for moisture effects in coupled land-atmosphere simulations in the Karakoram. This additional level of complexity could thus be investigated in further studies for hydrological applications.

The debris component of the model implemented here again draws on the algorithms for surface energy balance and heat conduction at the core of FSM. Surface properties, including albedo and roughness length, are adjusted where debris is at the surface following Collier et al. (2015). As per the "dry" debris formulation in Collier et al. (2014), the latent heat flux at the surface is set to zero if debris is the uppermost model layer. Heat conduction between any snow overlying debris is accounted for analogous to the case of snow over soil, as are heat fluxes between the base of the debris and the upper glacier layer, albeit with appropriate specific heat capacity, density and thermal conductivity values for each layer. Energy from heat conduction through debris is the only source of energy driving glacier melt where debris is present. The debris is discretised vertically into thin layers of 2 cm, increasing to maximum layer thicknesses of 10 cm for the lower parts of the thickest debris cover. While Collier et al. (2014, 2015) apply an estimated porosity function to help estimate the change in debris properties with depth, we follow the simpler approach of prescribing bulk properties that are constant with depth and in time, as per Reid and Brock (2010).

Similar to Collier et al. (2015), debris thickness was estimated based on distance along the glacier flowline from the upper limit of debris cover. As Collier et al. note, this appears to represent an improvement over parameterising debris thickness based on elevation alone. Collier et al. suggest that the linear relationship employed in their study may overestimate glacier thickness in the middle parts of the debris-covered ablation zone, while underestimating debris thickness towards glacier termini, although the latter issue may implicitly help to represent the ablation hotspots formed by supraglacial ice cliffs and ponds. To account slightly more for the nonlinear increase in debris thickness near glacier termini, we used an exponential function to estimate debris thickness based on distance along the flowline (up to a limit of 1 m thickness). The function is based on the debris thickness maps produced for the Baltoro glacier by Mihalcea et al. (2008) and Gibson et al. (2017), who in turn used field measurements, remote sensing and inverse modelling. The uncertainty here is of course large, but the strategy does represent a first estimate consistent with available information.

In line with most other models of sub-debris melt processes, the effect of very thin debris cover of up to 2-3 cm (or patchy cover) on accelerating melt relative to the clean ice case is not accounted for. Moreover, the role of supraglacial ice cliffs and ponds as "hotspots" of ablation is also not considered. In both cases, adequate representations for the larger scale simulations required to support water resources management are currently lacking. One option might be to use an effective debris thickness, which is lower than the actual thickness in heavy debris cover, although data to establish a possible relationship between effective and actual thickness are unavailable. The significance of these processes at the sub-basin and basin scales in the UIB certainly needs to be investigated, but for now we presume that the foremost effect of debris cover, especially thicker covers, is to suppress ablation (e.g. Hewitt 2014; Vincent et al. 2016).

Empirical Ablation Models

To permit comparisons with the energy balance approach to ablation modelling in FSM+, several alternative approaches to melt computation have been implemented. That these methods are incorporated as options within FSM+ means that they can utilise all of the same climate inputs, static inputs and parameters, as far as is relevant for any given formulation. For snow and clean ice, both temperature index (TI) and enhanced temperature index (ETI) models were implemented. The ETI method is described above, while the TI method takes the following simple form:

$$M = \begin{cases} TF \cdot T, & T > T_T \\ 0, & T \leq T_T \end{cases} \quad (3)$$

where the temperature factor (TF) is also referred to as a degree day factor (DDF). The TI and ETI models can also be run for glaciated cells with debris cover by adjusting the relevant parameters (referred to as TId and ETId methods, respectively). In addition, we implemented the new DETI parameterisation for sub-debris melt described earlier (Section 5.2.4, equation 2). A range of parameter values have been tested for the empirical parameterisations. These ranges were determined from the literature, based on Hock (2003) for the TI method, as well as Finger et al. (2011) and Ragettli et al. (2013, 2015) for the ETI method. Less testing appears to have been undertaken for parameters modified for debris cover, such that we focus on the benchmarks provided by Lutz et al. (2016a) (TId), Ragettli et al. (2013) (ETId), and Ragettli et al. (2015) and Carenzo et al. (2016) (DETI).

Mass Redistribution

With the emphasis of this work on ablation model inter-comparison, only simple parameterisations of snow and ice mass redistribution have been incorporated into FSM+ so far. Similar to TOPKAPI-ETH, these are the SnowSlide algorithms for gravitational redistribution of snow (Bernhardt and Schulz 2010), as well as the Δh-parameterisation of glacier flow. Further work could examine in more detail the implications of these model representations. For snow redistribution, it would be beneficial to examine the combined effects of wind and gravity on snow distributions and the high elevation water balance. Strasser et al. (2008) demonstrated the substantial influence of wind-related processes on sublimation in high, exposed locations in an Alpine context, although Bernhardt et al. (2012) showed that this is substantially reduced if a parameterisation of avalanching is included. Incorporating SnowSlide in FSM+ thus likely helps to reduce the error from omission of wind redistribution effects, if not completely so.

For glacier flow, developments could include an approach related to the shallow ice approximation (SIA), similar to Forsythe (2012), Immerzeel et al. (2012a) or Plummer and Phillips (2003). Some of the assumptions underpinning this approach are unlikely to hold ubiquitously in the steep mountains of the UIB, but this level of complexity may nevertheless be reasonable to examine long-term glacier evolution in climate stabilisation scenario exercises. In such cases, the geometry of glaciers in an approximate dynamic equilibrium may be strongly controlled by mass balance, even if the SIA transient response would differ from higher order ice modelling. An appropriate flux limiter to guarantee mass conservation would also need to be implemented (Jarosch et al. 2013). This would help to answer questions about whether the Δh-parameterisation, with parameter transfer from the literature, deviates significantly from approaches based on simplified ice flow physics in applications like climate change projections.

Hydrology

With the runoff response of much of the UIB dominated by its seasonal and perennial cryosphere, we take fairly simple approaches to inclusion of some of the additional processes in the hydrological cycle at this stage. Potential evapotranspiration for cells without any snow or glacier cover in FSM+ is estimated using the FAO56 Penman-Monteith approach for reference evapotranspiration (ET0), which is multiplied by crop factors that vary monthly and as a function of land use. Along with surface and subsurface flow routing, actual

evapotranspiration is calculated following the open source Python implementation of TOPKAPI (PyTOPKAPI - Sinclair and Pegram 2013). The parameters were assigned to be equivalent to those used in TOPKAPI-ETH (Section 5.2.4). Future work could clearly explore the potential for more physically based models to represent the surface and subsurface hydrological flows in the UIB. These models may be quite successful even in the absence of detailed information on catchment properties in the UIB, potentially after some calibration of the most sensitive parameters. With the emphasis on ablation processes in this chapter, we leave detailed analysis of these aspects of UIB hydrology for further work.

5.3 Results and Discussion

This section begins by analysing the TOPKAPI-ETH benchmark. The model's parameter sensitivity is then explored, which is followed by an examination of parameter optimality compared with re-calibration using a multi-objective optimisation algorithm. The analysis then moves to FSM+, beginning with a comparison of the underlying FSM snow model configuration with empirical alternatives. Examining the relationship between surface energy balance terms and melt rates subsequently leads to some insights into systematic differences between approaches. Thereafter, the focus moves to glacier ablation. Following appraisal of the glacier component developed for FSM+, the clean and debris-covered components of glacier ablation are assessed in turn relative to empirical alternatives. The implications of differing climate sensitivities in glacier ablation approaches are then considered in a simple temperature perturbation test. Finally, the analysis compares the catchment-scale performance and water balance components of TOPKAPI-ETH and FSM+.

5.3.1 TOPKAPI-ETH Benchmark

Figure 5.1 compares observed annual cycles of river flow with simulations using the TOPKAPI-ETH benchmark described in Section 5.2.4. Overall this shows that the magnitude and timing of flows are reasonably reproduced for the Astore and Hunza sub-basins by the combination of inputs and parameters employed. Figure 5.1 also shows that the TOPKAPI-ETH models capture the range of inter-annual flow variability. Strong agreement is evident for the snow-dominated Astore sub-basin (Figure 5.1a), albeit with a possible slight high bias in the July peak flow. For the heavily glaciated Hunza sub-basin, flows in TOPKAPI-ETH appear to rise and peak a little earlier than in observations (Figure 5.1b). This could reflect some deficiency in the volume of simulated glacier melt, which is considered to be a key component of river flows during this month (Mukhopadhyay and Khan 2015).

Underestimation of melt is one possibility, but alternatively this could reflect some low bias in glacier volume following initialisation and/or warm-up. That the simulated Hunza flows peak in July, rather than being sustained through August, could also indicate some remnant bias in winter/spring snowfall amount. Specifically, surplus seasonal snow may remain on the ground into July, a time of high energy availability for melt.

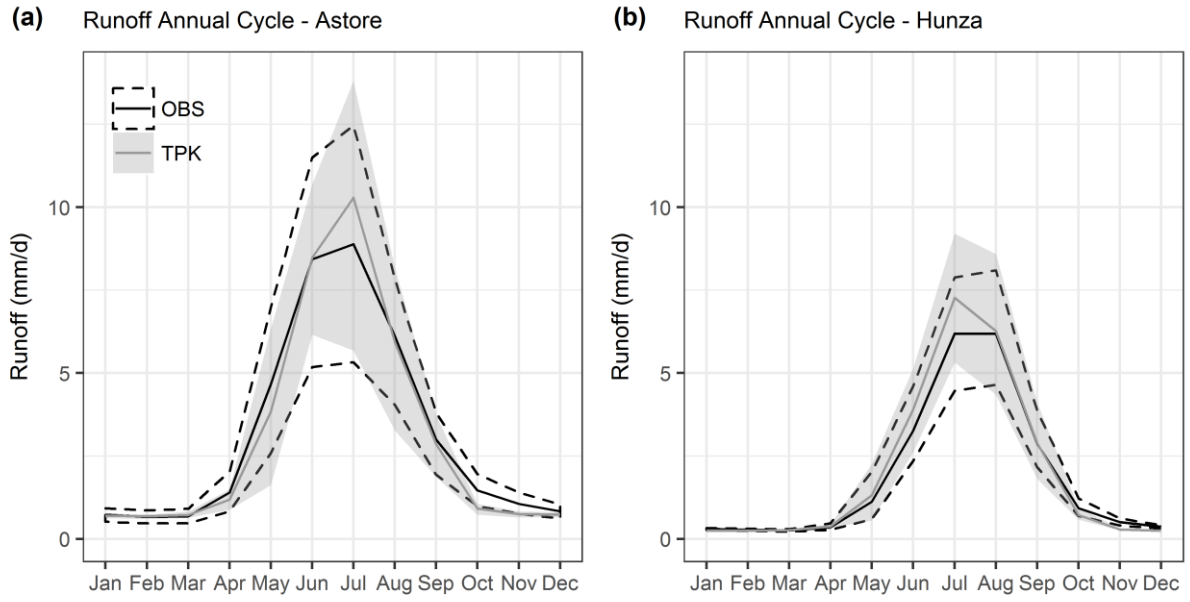


Figure 5.1 – Observed and simulated (TOPKAPI-ETH baseline) annual cycles of rivers flows in the (a) Astore and (b) Hunza sub-basins. OBS refers to observations and TPK refers to TOPKAPI-ETH.

To confirm that the annual cycles in Figure 5.1 are associated with plausible shorter term variability, time series of daily flows are shown in Figure 5.2. Generally this suggests that TOPKAPI-ETH captures sub-seasonal variability reasonably well, with the alignment of observed and simulated variability adding some further confidence that the climatic inputs are reasonable and the model responses plausible. Some differences compared with observations are of course apparent, as might be reasonably expected in such data-sparse high mountain catchments. Interestingly, the benchmark configuration of TOPKAPI-ETH appears to show a comparatively peaky response for the Astore sub-basin relative to observations. With the exception of some peaks in summer 2004, the Hunza sub-basin does not appear to be so starkly afflicted by a similar issue. This begins to raise the question of whether the benchmark parameters are globally optimal across both snow- and glacier-dominated hydrological regimes in the UIB, as the Astore simulation in effect represents a blind validation of the configuration used by Ragettli et al. (2013) for the Hunza.

Figure 5.2a also suggests that the rising limb of the river flow hydrograph in spring tends to be delayed in the Astore simulation. This could reflect some residual bias in temperatures, a degree of underestimation of snowfall at lower elevations, or some issue with the parameters controlling melt and routing.

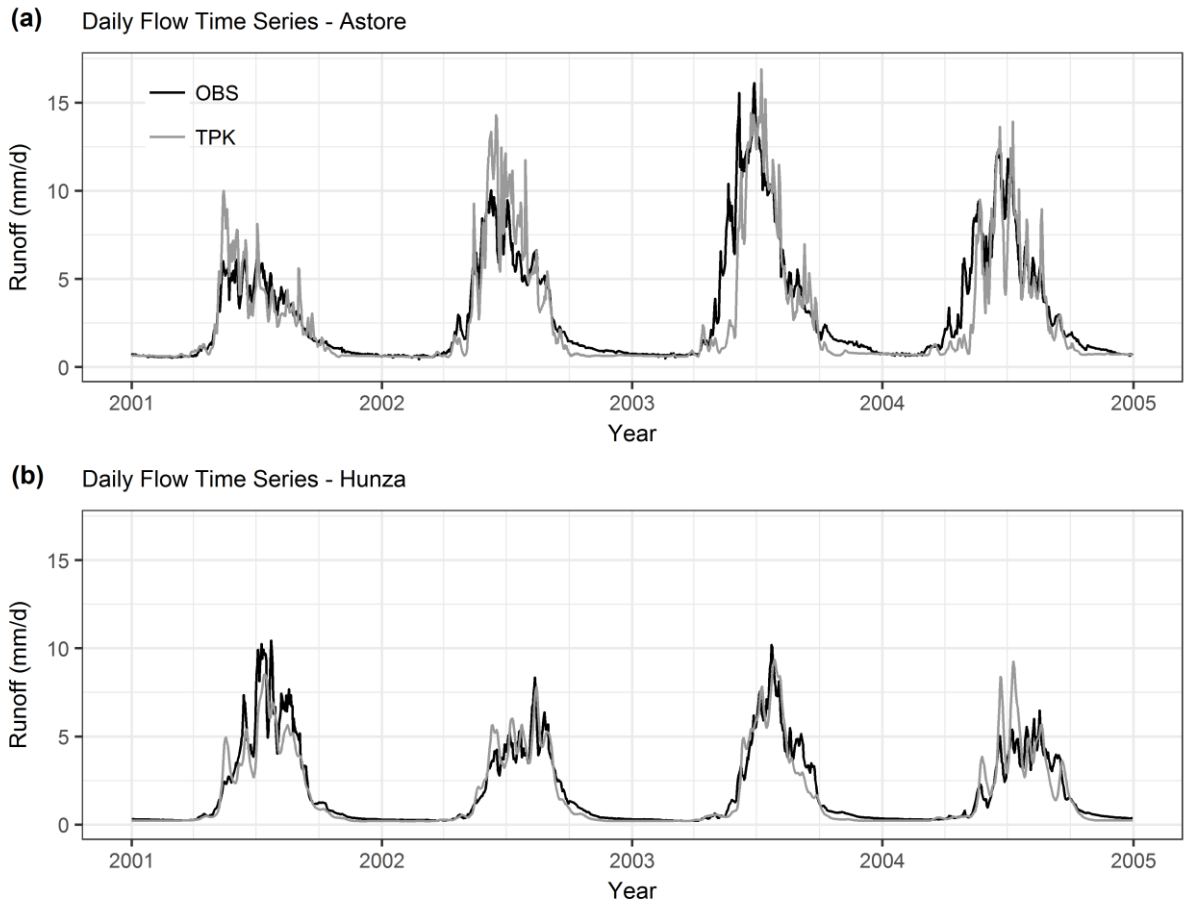


Figure 5.2 – Observed and simulated (TOPKAPI-ETH baseline) daily rivers flow time series in the (a) Astore and (b) Hunza sub-basins.

As a means of assessing runoff generation processes internal to the catchments, Figure 5.3 compares the annual cycle of catchment snow-covered area (SCA) variation in TOPKAPI-ETH with MODIS SCA. This helps to identify whether accurate catchment runoff is obtained for the right reasons, confirming the physical plausibility of the internal catchment response. Figure 5.3 shows that there is generally reasonable agreement between the simulated and remotely sensed annual cycles. Overall, this suggests that the cycles of mass and energy inputs to the sub-basins are reasonably well approximated, although there are some differences relative to MODIS. For the Astore sub-basin, the rate of SCA disappearance is initially a little faster than in MODIS, with the TOPKAPI-ETH annual minimum remaining slightly higher. Inter-annual variation is mostly consistent between TOPKAPI-ETH and

MODIS for Astore, with the exception of a discrepancy in November, most likely due to a precipitation timing issue in the driving dataset.

In general, TOPKAPI-ETH also captures the fundamental difference in SCA cycles between the snow-dominated Astore and glacier-dominated Hunza sub-basins. Most notably, the smaller amplitude of the annual cycle in the Hunza is captured, along with the slower rate of SCA decline in spring/summer and the higher annual minimum SCA. These features of the simulated Hunza SCA cycle are encouraging, but there are also some disagreements with respect to MODIS. TOPKAPI-ETH simulates a higher winter snow cover extent than that suggested by MODIS, as well as substantially less inter-annual variability in winter SCA. The spring/summer SCA reduction is also comparatively rapid in TOPKAPI-ETH, leading to an annual minimum below its MODIS counterpart. Some of the differences could relate to difficulties in directly comparing model output with MODIS, for example due to the need to define thresholds for snow cover classification. However, Figure 5.3 suggests that the differences may be reasonably systematic, reflecting some level of input or model limitations.

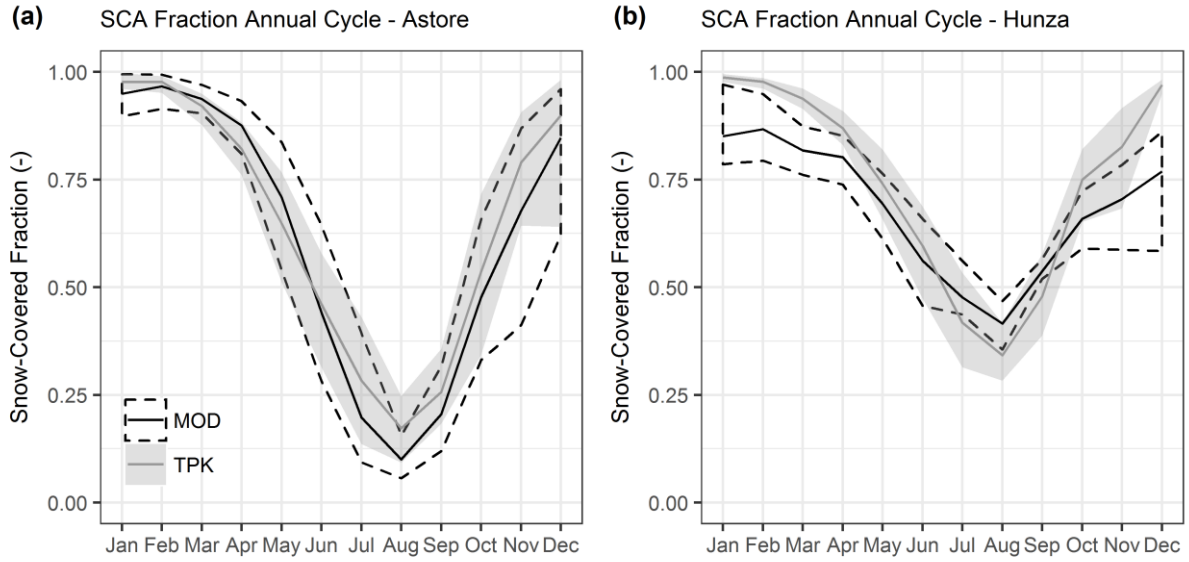


Figure 5.3 – Remote sensing (MODIS) and simulated (TOPKAPI-ETH baseline) annual cycles of snow-covered area (SCA) in the (a) Astore and (b) Hunza sub- basins. MOD stands for MODIS and TPK denotes TOPKAPI-ETH.

To examine the skill of TOPKAPI-ETH forced by the HAR in capturing the structure as well as the range of inter-annual variability, Figure 5.4 compares simulated and MODIS monthly SCA anomaly time series. The anomaly for each month in the series is calculated as the difference in SCA compared with the period mean for that month. Figure 5.4 suggests that the

sequencing of anomalies is fairly well reproduced by the model. Particularly for the Astore sub-basin, the alignment of positive and negative SCA anomalies in Figure 5.4 is good, with the magnitude of anomalies often well simulated too. The agreement for the Hunza sub-basin appears slightly lesser, although the series do appear to generally track.

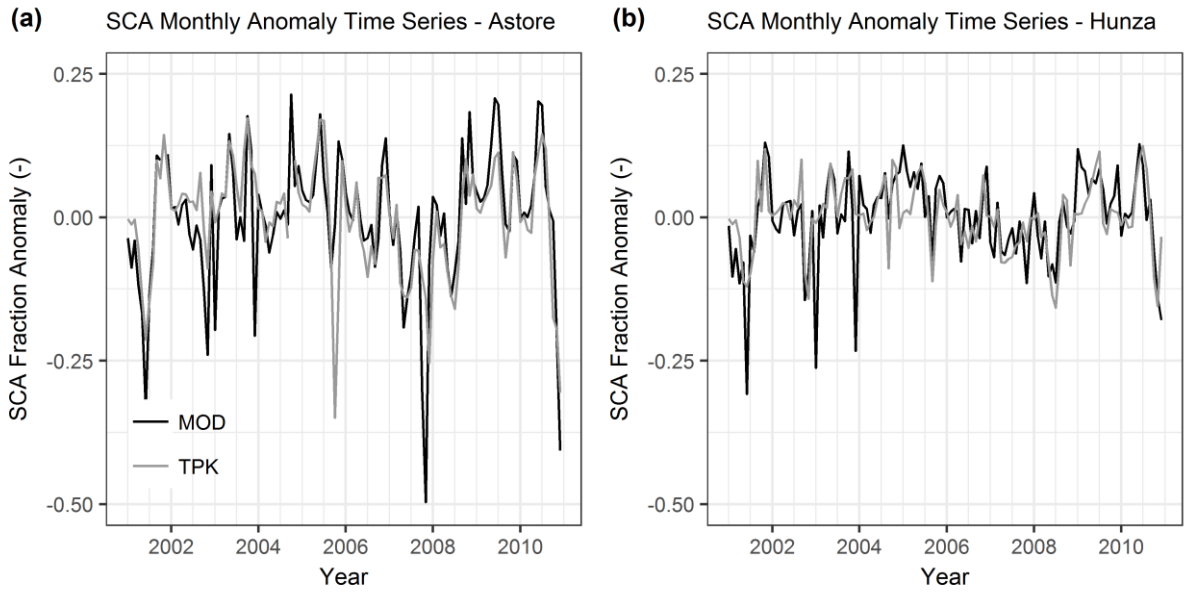


Figure 5.4 – Remote sensing (MODIS) and simulated (TOPKAPI-ETH baseline) time series of snow-covered area (SCA) anomalies in the (a) Astore and (b) Hunza sub-basins.

5.3.2 TOPKAPI-ETH Sensitivity Analysis

While the baseline TOPKAPI-ETH model run thus provides generally reasonable performance, this configuration is only one realisation of an array of uncertain input and parameter possibilities. To explore this further, the Morris method (Morris 1991) was selected as a practical approach for screening-level sensitivity analysis, focusing on the Astore sub-basin. The Morris method is essentially a variant of a one-at-a-time local sensitivity method and begins by defining a possible range of values for each parameter. For each of the r sets of starting values of these parameters, an initial model run is carried out. Then each of the k parameters is adjusted in turn from its starting value. Within any one of the r sets, once a parameter has been adjusted from its starting value, it retains its new value until all parameters in the set have been altered. This means that $r(k + 1)$ runs are required. The method thus exhibits comparatively low computational requirements. In an inter-comparison of sensitivity analysis methods, Gan et al. (2014) showed the Morris method to be efficient and generally adept at distinguishing between sensitive and insensitive parameters in a

hydrological model, although they note that it is not as robust as some more intensive alternatives.

Absolute differences in a specified model output compared with the starting values are calculated for each run and then aggregated separately for each parameter. Aggregation is based on the mean (μ^*) and standard deviation (σ). The position of parameters along the μ^* (x) and σ (y) axes (see Figure 5.5) represents whether the model is consistently sensitive to a parameter/input (mu direction) or if its sensitivity varies depending on other parameters (sigma direction).

In contrast to other TOPKAPI-ETH sensitivity analyses restricted to cryosphere-related parameters, we expose all of the adjustable parameters in our analysis. This includes those parameters related to evapotranspiration, soil/groundwater and routing, as well as parameters governing cryospheric processes. Parameter values are based on ranges derived from the literature (Ragettli et al. 2013). To examine uncertainty in climate inputs, as well as the its importance relative to model parameters, the elevation gradients used to help disaggregate the HAR precipitation and temperature field were based on a monthly gradient climatology, inferred from the HAR, and allowed to vary as parameters in the sensitivity test. A presumed bias in cloud transmissivity fields was also permitted. The ranges for the precipitation and temperature gradient tests were set based on the ranges tested for calibration in Ragettli et al. (2013).

Figure 5.5 suggests that the most influential parameters in the TOPKAPI-ETH baseline model for the Astore sub-basin fall into several groups. Precipitation gradients and temperature lapse rates used to specify spatiotemporal climate fields are consistently important, along with any potential bias in cloud transmissivity. This is consistent with precipitation in the preceding winter/spring strongly determining summer runoff in the snow-dominated Astore sub-basin, as well as the high sensitivity of TOPKAPI-ETH to precipitation and temperature gradient variability demonstrated elsewhere in the Himalayan arc (Immerzeel et al. 2014). Intuitively, the shortwave and temperature factors of the ETI melt model are also strongly influential. In line with this, the two parameters governing the albedo parameterisation in the TOPKAPI-ETH ETI implementation also show importance. There is strong similarity between Figure 5.5a and Figure 5.5b in terms of the parameters implicated as predominant in modelled runoff and SCA responses, respectively, highlighting the strong mutual information in these datasets. The main difference between the two cases is that flow response does show some sensitivity

to some of the subsurface hydraulic properties, whereas the SCA-based analysis flags up precipitation phase partitioning as more significant. These contrasts arise due to the information provided by different evaluation criteria, i.e. integrated catchment mass input/output compared with internal snow cover state.

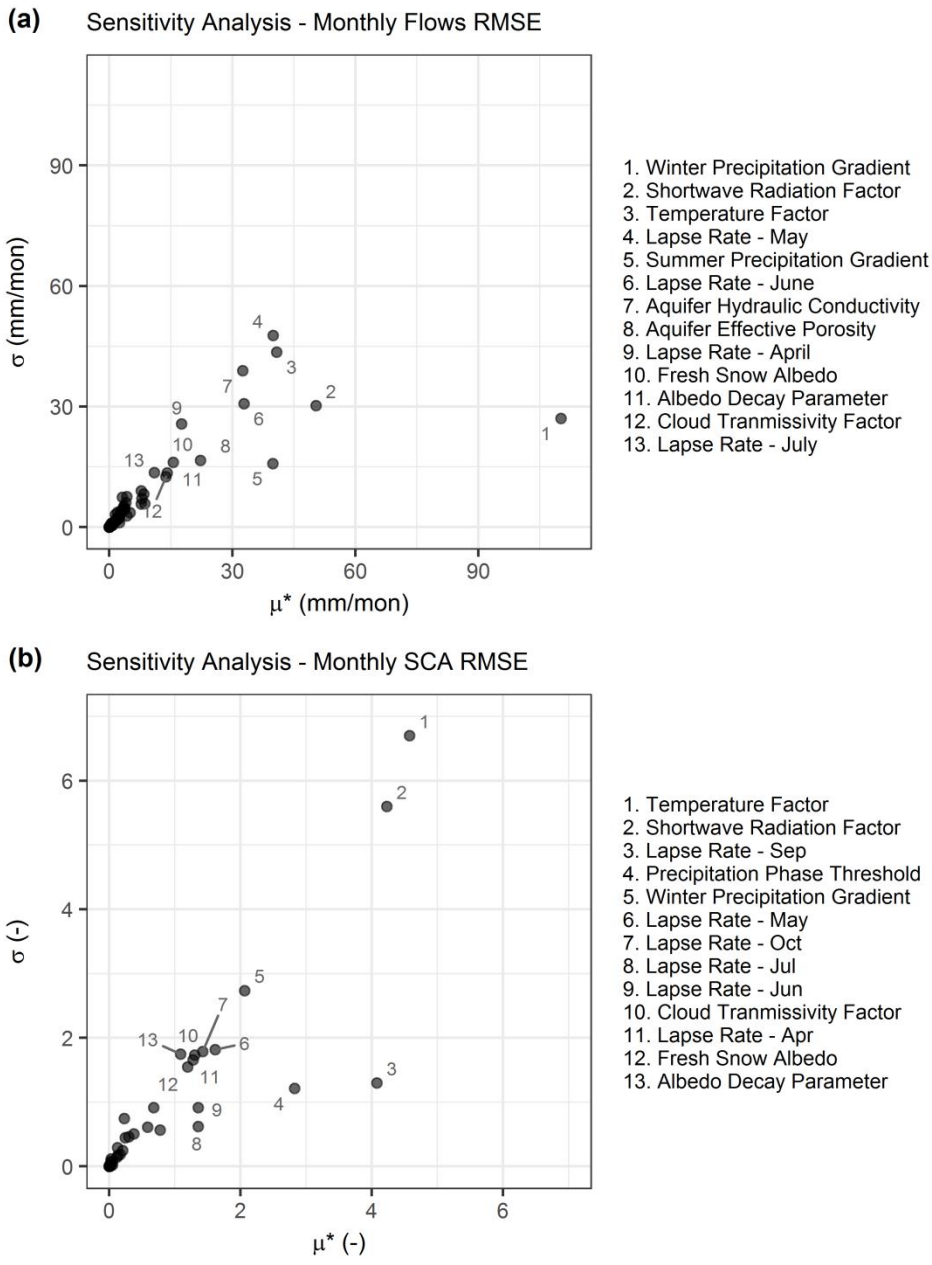


Figure 5.5 – Sensitivity analysis based on the Morris method for (a) flows and (b) SCA for the Astore sub-basin.

These findings are largely in line with expectations, but we note that the time period of the sensitivity test is relatively short (restricted to the length of the HAR dataset). Although reasonably short simulations help to facilitate sensitivity tests for models with high

dimensional parameter spaces, it would be interesting to investigate the relationship between parameters emerging as sensitive and the timescale of response in more detail. For example, glacier mass balance over longer periods could be more sensitive to parameters controlling snow redistribution at high elevations than may be inferred from short simulations alone. Moreover, while the method is primarily suited for qualitative, rank-based analysis, it is interesting to note the overall relationship between *mu** and *sigma*. Thus, even for parameters identified as consistently important in their own right (high *mu**), there appears to be substantial dependence on the values of other parameters. This may reflect the relatively high-dimensional nature of the problem and the non-uniqueness of model input/parameter combinations in relation to achieving ostensibly adequate model behaviour with respect to observations.

5.3.3 TOKPAPI-ETH Multi-Objective Optimisation

The sensitivity screening analysis suggests that the ETI and albedo parameters are of critical importance, along with uncertainty in the key climate forcings. To explore further whether the ETI and albedo parameters in the benchmark are approximately optimal, we evaluated other possibilities using multi-objective optimisation. The purpose of this is to see whether the benchmark values for the most critical parameters are one of the better-performing sets, given the HAR climate forcing and other aspects of model configuration. The three objective functions specified in this application were minimisation of errors in monthly river flows, SCA and glacial mass balance. The latter criterion assumes an approximately neutral glacier mass balance over the simulation period, in line with the literature for this area (Hewitt 2005; Bolch et al. 2012; Gardelle et al. 2013; Kääb et al. 2012; Gardner et al. 2013; Bolch et al. 2017; Zhou et al. 2017), as more specific, sub-catchment scale glacier mass balance data are unavailable. Nevertheless, recent high resolution estimates for a period largely coincident with the HAR record suggest this assumption may be a reasonable starting point (Brun et al. 2017). The widely applied non-dominated sorting genetic algorithm NSGA-II was used (Deb et al. 2002), while climate inputs were fixed following the method described in Section 5.2.1, such that only ETI and albedo model parameters were permitted to vary within the same ranges from the literature used in the sensitivity analysis above.

Using multiple objectives means that an individual solution is not generally present; rather, solutions lie on a Pareto front. Figure 5.6 shows two-dimensional visualisations of the trade-offs between objective functions, and how the form of the Pareto front varies depending on

the dimensions of the objective function under consideration. In Figure 5.6a, the lower left corner, representing the best performing parameter sets, actually converge strongly towards a point. Better performance in SCA simulation is thus strongly linked with better river flow performance. These two datasets are thus highly complementary, ostensibly sharing a lot of mutual information. In contrast, Figure 5.6b illustrates that trade-offs between glacier mass balance simulation and river flow simulation errors induce a clear Pareto front. The smallest errors in river simulation do not coincide with the smallest errors in presumed glacier mass balance. This may be partly to do with the approximation of the glacier mass balance target, in the absence of sufficient data. Alternatively, it may reflect unavoidable trade-offs in multi-objective optimisation, which could be compounded if input or structural model errors are present, although the curve of the Pareto front is quite steep, such that a relatively large apparent performance gain in glacier mass balance may be obtained with a relatively small decrease in river flow performance.

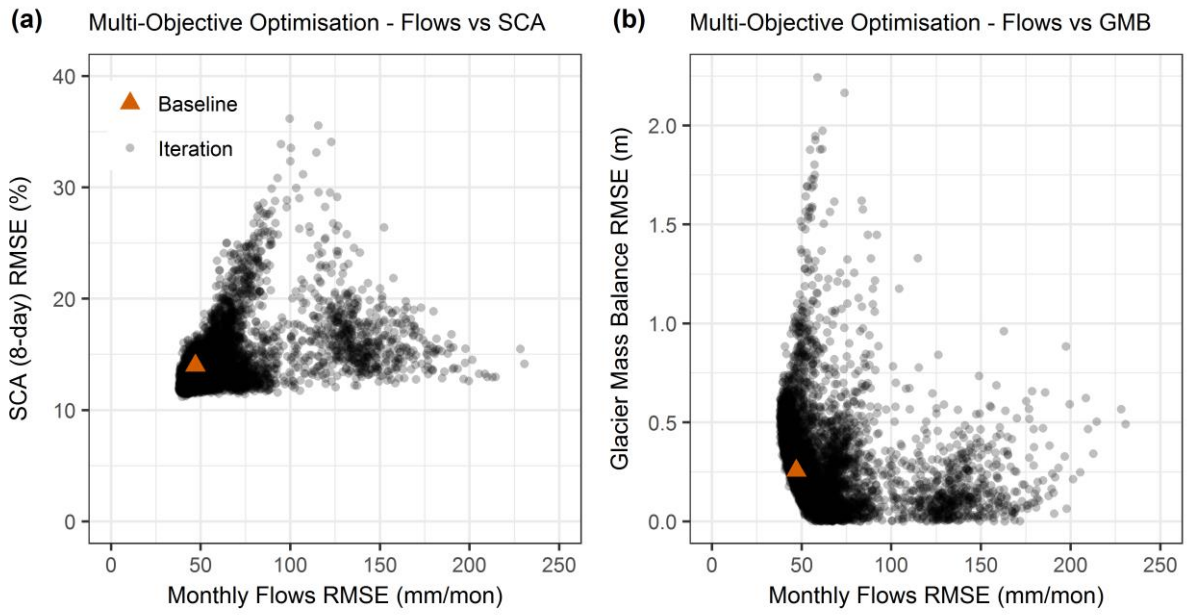


Figure 5.6 – Trade-offs in multi-objective evolutionary optimisation for monthly flow and SCA errors (a), as well as monthly flow and overall glacier mass balance (b).

Figure 5.6 also shows that the baseline parameter set (Section 5.2.4) falls relatively close to the fronts in both panels. This suggests that the parameter set is a reasonable choice in relation to the input data, which is in turn considered a fair approximation in the absence of more data. Yet, in neither case does the baseline TOPKAPI-ETH configuration lie on the Pareto front. As such, it seems that the multi-objective optimisation approach here can guide some parameter refinement relative to the simpler approach taken by Ragettli et al. (2013) that underlies the

baseline configuration. Despite this, the baseline model configuration does appear to show reasonably good performance with respect to available data sources, which helps to provide some further, partial validation of the ETI approach, as well as the climate and static inputs described in Sections 5.2.1 and 5.2.2. Therefore we consider it appropriate to take this intersection of model structure and input datasets forward for deconstruction in the remainder of this section.

5.3.4 Snow Ablation – Baseline Comparison

Sections 5.3.1-5.3.3 suggest that the TOPKAPI-ETH configuration based on Ragettli et al. (2013) provides reasonable performance for the Astore and Hunza sub-basins of the UIB, with respect to observed flows and MODIS SCA. In this section, we compare the ETI approach to snowmelt modelling from TOPKAPI-ETH with the energy balance approach adopted in FSM. As described in Section 5.2.5, we also incorporate comparisons with the TI method, using TI and ETI models implemented within the FSM+ framework.

Figure 5.7 compares catchment-scale daily snowpack runoff calculated using the TI, ETI and FSM models for the snow-dominated Astore sub-basin for three contrasting melt seasons. The aim of this initial comparison is to provide an indication of how different modelled melt generation might be when using fairly typical parameter values. As noted in Section 5.2.5, the comparisons are provided for two members of the FSM+ ensemble: one with the stability option switched off (ST-0) and one with it switched on (ST-1). The snow degree day factor in the baseline TI model (5 mm °C d^{-1}) was based on previous studies (Immerzeel et al. 2015; Lutz et al. 2016a), whereas the ETI parameters were taken directly from the benchmark TOPKAPI-ETH setup. Observed river flows at the catchment outlet are plotted for reference; however, they are not directly comparable with the model series, which show runoff released from the base of the catchment snowpack (i.e. melt released after any delay due to vertical flow and melt/refreezing cycles – referred to as snowpack runoff hereafter). The observed flows serve as a useful reference regarding the magnitude of melt and runoff, as well as the timing of the annual cycle of flows and intra-annual variability. We elected to evaluate the modelled series in this way to avoid some of the complicating influences of routing.

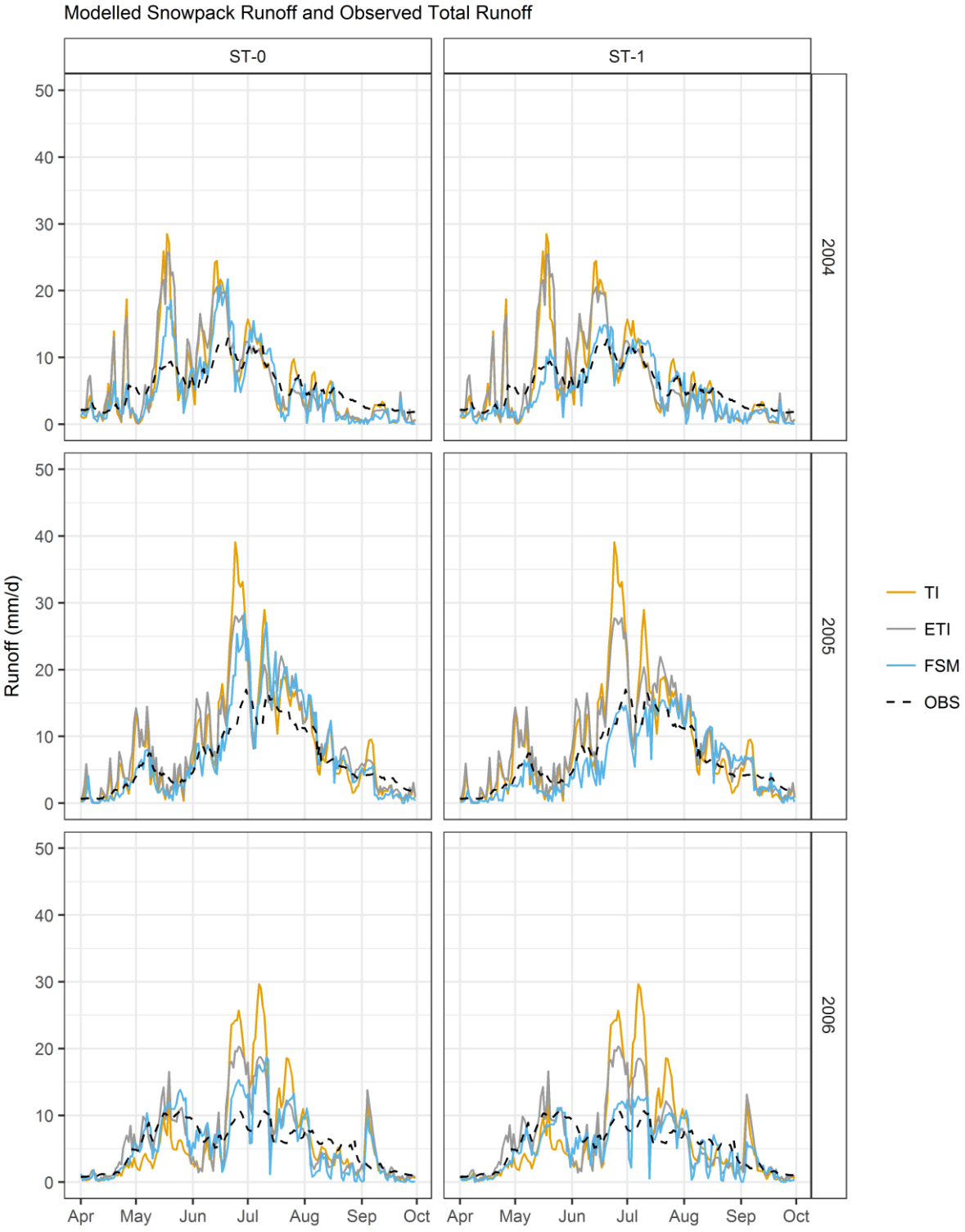


Figure 5.7 – Comparison of catchment-scale snowpack runoff calculated using the FSM energy balance model (blue) and (TI) a temperature index approach (orange) and enhanced temperature index (ETI) (grey) approach for three example years. FSM configurations with stability adjustment off (ST-0) and on (ST-1) are applied. Observed runoff is also shown (black dashed line).

A number of points are evident from Figure 5.7. Firstly, the general tendency for FSM with ST-1 to exhibit a less peaky, more subdued melt response than ST-0 is apparent. This pertains

to the discussion in Chapter 4 about the strength of the required stability correction for turbulent fluxes, as well as the degree to which a relatively high sensible heat flux may be realistic or potentially compensate for sub-grid variability or topographically modulated influences on the surface energy balance, such as terrain enhancement of incoming longwave radiation. In general, Figure 5.7 shows greater correspondence between the TI/ETI methods and FSM when applying the ST-0 option. This could suggest that, at least for the baseline parameters here derived from the literature, the two empirical methods are more akin to an energy balance approach in which stability adjustment is treated at its limit, i.e. completely unadjusted. The generality of this point is explored shortly with respect to parameter sensitivity tests. Yet, even compared with ST-0, the TI and ETI models appear to show relative volatility, such that their day-to-day variability is fairly pronounced relative to both FSM cases.

In general, the TI and ETI models show more snowpack runoff early in the melt season or, at the very least, more spikes in the early response. This reflects the notable role played by cold content and particularly melt/refreezing cycles early in the season. Without some incorporation of this the TI/ETI models show flashy melt generation patterns on warm days. However, the magnitude and frequency of these relatively early melt pulses in the TI and ETI models can be quite variable. Moreover, while the TI and ETI melt patterns are strongly in agreement in 2004 and 2005, much more divergence is apparent in 2006. Specifically, the TI response is comparatively delayed, while the ETI matches more closely with the FSM series. This suggests that relatively cold temperatures preclude significant runoff generation in the TI model in the early part of the 2006 melt season. That more substantial melt is sustained by the ETI model suggests that the temperatures must be large enough to permit melt, but that this must be largely driven by the net shortwave radiation term in the ETI baseline. In this case there appears to be greater agreement between the ETI and FSM models.

5.3.5 Empirical Snow Model Parameter Uncertainty

While some differences in snowpack runoff responses are evident in Figure 5.7, the baseline TI and ETI models in that comparison use parameters assigned a priori. This raises the question of how the comparison of TI/ETI models and FSM might differ with alternative parameter choices. To examine this, parameters were varied systematically for the TI and ETI models in separate runs (i.e. each parameter was held constant during a run). To some extent this replicates how the different types of models are typically used in the Himalayan region,

namely without accounting for spatial and temporal parameter variation but with calibration to attain effective or average parameters (e.g. Bocchiola et al. 2011; Rees and Collins 2006; Immerzeel et al. 2009, 2012a, 2013; Soncini et al. 2014; Ragettli et al. 2013, 2015; Immerzeel et al. 2014; Minora et al. 2015). More complicated EB approaches, such as FSM, tend to be used with minimal calibration in distributed applications such as this, with as many parameters as possible set based on reference values. This is in part due to the higher complexity and run times of such models, which makes exploration of their high-dimensional parameter space more challenging. Although more of the parameters in FSM may be physically measureable, this is often not completely the case for practical reasons or where simplified parameterisations of some processes need to be adopted. A comprehensive parameter uncertainty analysis of the FSM configurations may be useful in further work.

TI – Catchment Scale

Figure 5.8 shows the mean differences in runoff between the TI and FSM models as a function of selected degree day factors (DDFs). Differences were calculated as TI minus FSM. The range of DDFs tested was based on the collation of values reported for snow in different studies in Hock (2003) (see Section 5.2.5). Figure 5.8 indicates that, for a given DDF, there may be notable variation in the correspondence between the two types of models in different months. In April, the choice of DDF has a relatively small impact on mean runoff differences in absolute terms, owing to the lower snowpack runoff in the energy-constrained ablation conditions at this time of year. Nevertheless, the variation appears significant in relative terms. A DDF of between 4 and 6 mm/d/°C is most consistent with FSM for ST-0, although a slightly lower DDF may be required to mimic ST-1. Relative to April, higher DDFs tend to be more consistent with FSM in May. The relationship between mean runoff difference and DDF steepens, while inter-annual variability increases substantially. Indeed, for most of the DDFs tested, the differences span both positive and negative values in different years. One possible reason for the higher inferred DDFs in May is that rising net shortwave radiation provides a key driver for melt that is essentially unaccounted for in the TI model. The omission of melt/refreezing cycles and their effect on snowpack ripening may also be important in permitting melt at comparatively low positive air temperatures. Thus, while air temperatures exceed the critical threshold and permit melt, higher DDFs appear to be needed to match FSM snowpack runoff rates. This applies in both ST-0 and ST-1 FSM cases, which are similar but offset.

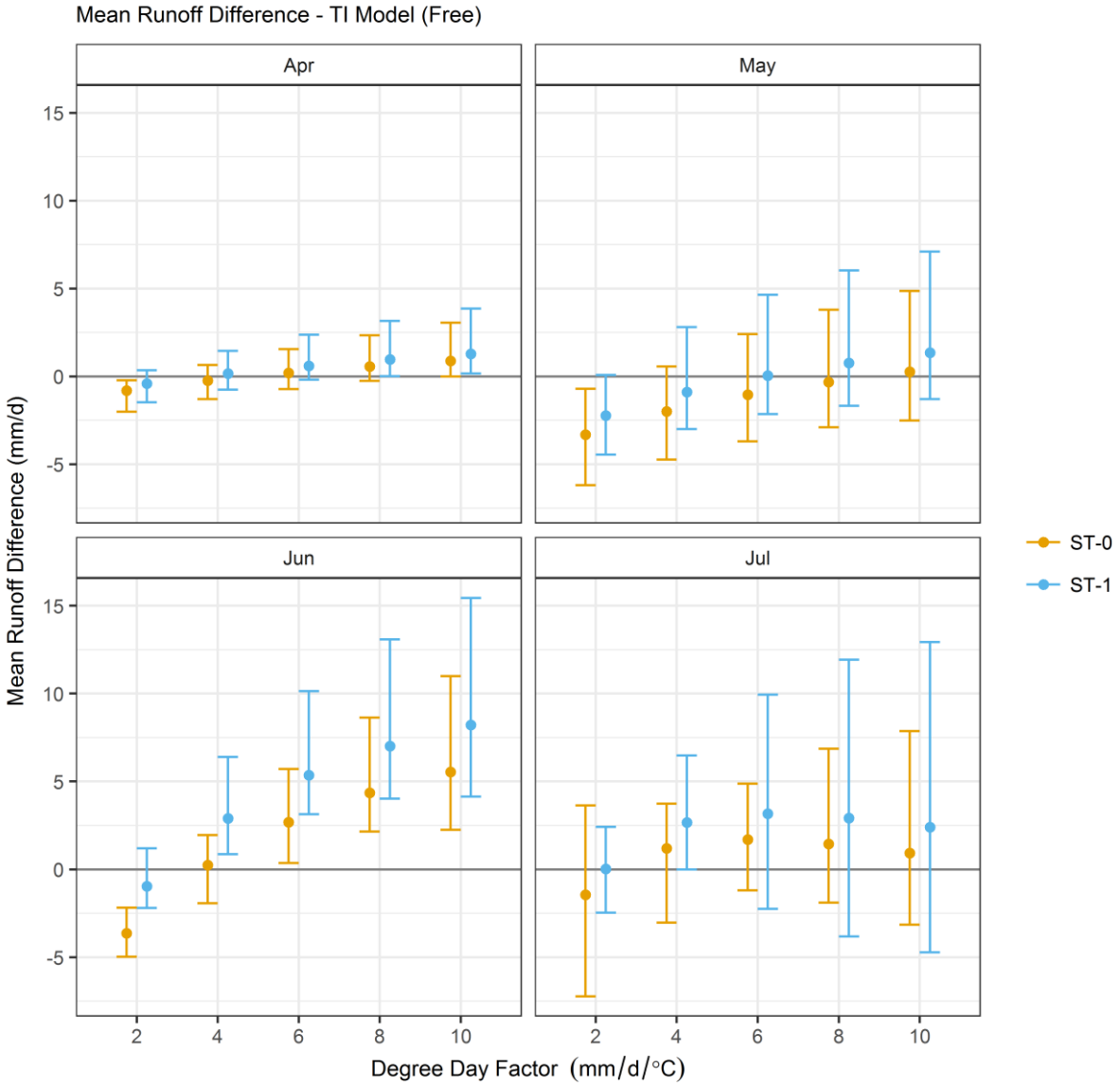


Figure 5.8 – Monthly mean differences (TI minus FSM) in snow melt between FSM and TI models for selected months. Both stability options are shown for FSM, with off (ST-0) and on (ST-1) in orange and blue respectively. Points show mean difference for all years of simulation, while lines show inter-annual ranges. Horizontal lines show the 10th and 90th percentiles and mean of observed runoff, which is for reference only (as model results are only for snow melt component of total flows).

The DDFs that might be required to mimic the FSM outputs in May are no longer appropriate in June. Rather, DDFs of 4 mm/d/°C or lower lead to more agreement between the models. As the higher DDFs are associated with higher melt in June, it is reasonable to consider the month as essentially energy- rather than mass-constrained in terms of ablation. Therefore Figure 5.8 suggests that for the FSM model there is a shift in the relationship between air temperature and snowpack runoff generation as the melt season progresses into June. As the melt season progresses, it could be that release of meltwater from the deeper snowpack at

higher elevations becomes more delayed relative to the shallower snowpack at lower elevations. This is explored further below.

By July a situation develops whereby the highest mean runoff differences are associated with moderate DDFs. The smallest DDF tested still leads to negative differences on average, which suggests that they are insufficient to reproduce the rates of melt in FSM at this time of year. The largest DDF tested gives positive differences, but these may be smaller than those induced by more moderate DDFs. This likely arises because the higher DDFs lead to more melt earlier in the season, limiting the mass available for ablation earlier than in TI realisations applying lower DDFs. The inter-annual variability in difference tends to increase with DDF in July, especially for ST-1. This is again to do with the climate conditions and rates of ablation in the preceding months. If a year has particularly large winter/spring precipitation and/or a relatively cold early melt season and/or a very warm July (i.e. permitting the freezing isotherm to get up particularly high), it is possible for large DDFs to induce relative high July melt, as more mass is available. Conversely, with climate conditions leaving less mass by July, high DDFs lead to particularly negative differences relative to FSM.

The analysis thus suggests that the DDFs most consistent with FSM snowpack runoff rates vary substantially throughout the melt season and on an inter-annual basis. This reflects the influence of variation in weather conditions. Yet, as the DDFs most consistent with FSM vary systematically within the melt season, rising and then falling, the results suggest that spatial and vertical patterns in the snow distribution may also be very important. This is explored below.

ETI – Catchment Scale

A similar analysis can also be conducted for the ETI model. Figure 5.9 again shows mean runoff differences relative to FSM, but this time for the ETI model and variation in its shortwave radiation factor (SRF) and temperature factor (TF). Beginning with the SRF of 0.004 m^2 mm W^{-1} hr^{-1}, which is at the lower end of the range reported by Ragettli et al. (2013), TF values towards the upper end of the range reported by Ragettli et al. (2013) tend to be required in May to match the snowpack runoff rates of the FSM models. However, in each case the range of inter-annual variability spans both positive and negative differences. Thereafter, lower or moderate TF values are more consistent with FSM in June. This is similar to the TI model, in that higher TFs (or DDFs) appear more consistent with FSM in

May compared with June, with both months being relatively unconstrained in terms of mass available for ablation. In July, there is some indication of stabilisation in the higher TFs, while the relationship between TF and inter-annual spread in mean runoff differences shows almost opposing signals for the ST-0 and ST-1 cases. While the inter-annual range decreases with higher TFs for ST-0, the inverse occurs for ST-1. This is likely a function of the ST-0 case becoming mass-constrained earlier. The similarity between the two higher TFs may suggest that the SRF term is sufficient to melt all of the snow, such that the sensitivity to TF reduces somewhat.

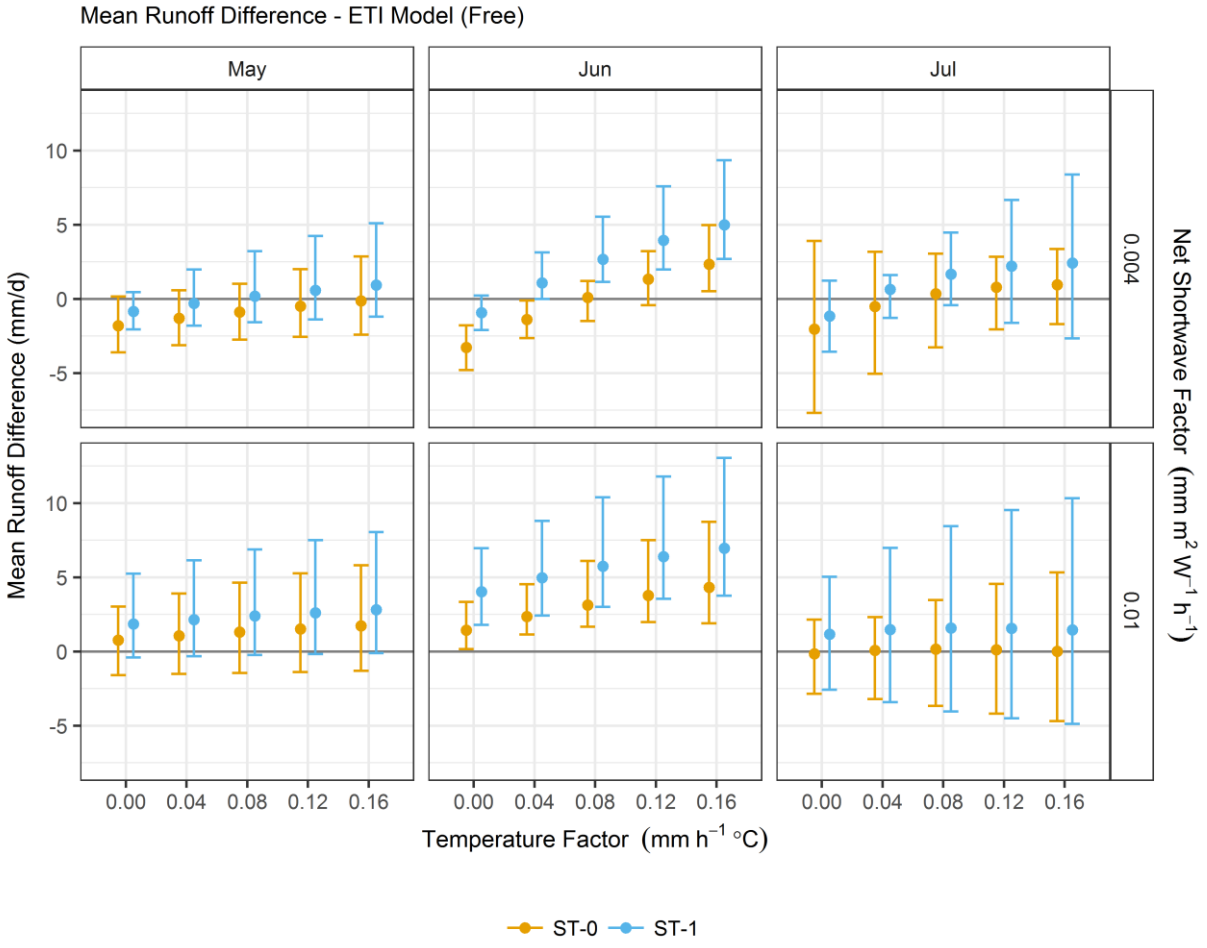


Figure 5.9 – Similar to Figure 5.8 but for ETI model. As the ETI model has two critical parameters, columns show the key months, while rows of the matrix are used for the shortwave radiation factor, with the temperature factor along the horizontal axis. The mean snowpack runoff difference is on the vertical axis, as per Figure 5.8.

With an SRF of 0.01 m^2 mm W^{-1} hr^{-1}, which is the baseline value taken from Ragettli et al. (2013) and approximately consistent with the coefficient needed to convert net shortwave radiation directly to a water equivalent depth, on average there is relatively little variation in

mean difference with different TFs for May. A slight positive relationship is evident. However, in both ST-0 and ST-1 cases, the ETI model leads to higher melt than FSM for all TF values, even if TF is set to zero, implying no temperature-driven component of melt. This suggests that the SRF parameter set at 0.01 m^2 mm W^{-1} hr^{-1} may in fact be relatively high, despite this being the baseline parameter and to some extent a physically interpretable value. Similarly, mean differences are high relative to FSM for all TFs during June, although in this case a stronger positive relationship between TF and snowpack runoff difference is evident. This fits with the generally abundant mass availability in June. By July, on average the ETI model is insensitive to the TF value, but the range of mean difference in runoff again increases with higher TFs. This suggests that sensitivity to the parameter at this time of year depends on the precipitation accumulation and ablation conditions, as well as contemporary temperatures. For ST-1, the ETI model still simulates higher snowpack runoff than FSM, owing to the overall slower rates of ablation in ST-1.

TI – Elevation Variation

An alternative approach to understand how TI/ETI parameters might vary to match an EB model is to inversely estimate the implied parameters based on the snowpack runoff simulated by FSM, in conjunction with the relevant climate fields. This approach is similar to that adopted by Raleigh and Clark (2014), in which the relationship between temperature and DDFs was examined for SNOTEL sites in the US. Raleigh and Clark used linear regression to identify the DDF at each site from the slope parameter of the relationship between cumulative positive degree days and cumulative snowmelt. The main difference in our application is that model outputs are used. The DDFs are estimated fortnightly for each grid cell by dividing the total snowpack runoff over the period by the cumulative positive degree days when snow was present.

Based on the possible significance of spatial variation inferred from Figure 5.8, these results are summarised by elevation bands in Figure 5.10. This shows that there is generally a characteristic annual cycle of DDF variation, whereby DDFs rise fairly rapidly with snowpack runoff onset and then later decline. The decline may reflect late-lying snow on north-facing slopes, but in the case of the 4500-5000 mASL elevation band, if anything there appears to be a stabilisation in the inferred DDF until melt ceases, at which point the DDF becomes undefined. The amplitude of the annual DDF cycle tends to be larger for lower elevation bands, in agreement with the results in Figure 5.8, which showed how high DDFs were

required relatively early in the melt season (e.g. May) to reproduce the FSM snowpack runoff rates at this time of year on average. However, the inter-annual variability in inversely estimated DDFs also tends to be larger at lower elevations, suggesting substantial variation. In general the patterns of intra-annual variation are similar for ST-0 and ST-1, with the main difference being the magnitude of inferred DDFs. Together, these results suggest that both spatial and temporal variation in DDFs would be needed if the EB approach of FSM were to be replicated, although the latter is of course simply a reference to a different, more process-based modelling approach rather than a truth.

Figure 5.10 – Annual cycles of degree day factors inversely estimated from FSM shown for different elevation bands. FSM stability options are shown in orange (ST-0) and blue (ST-1) respectively. The catchment-average degree day factors are shown in black. Lines denote the mean and shading shows the range of inter-annual variation.

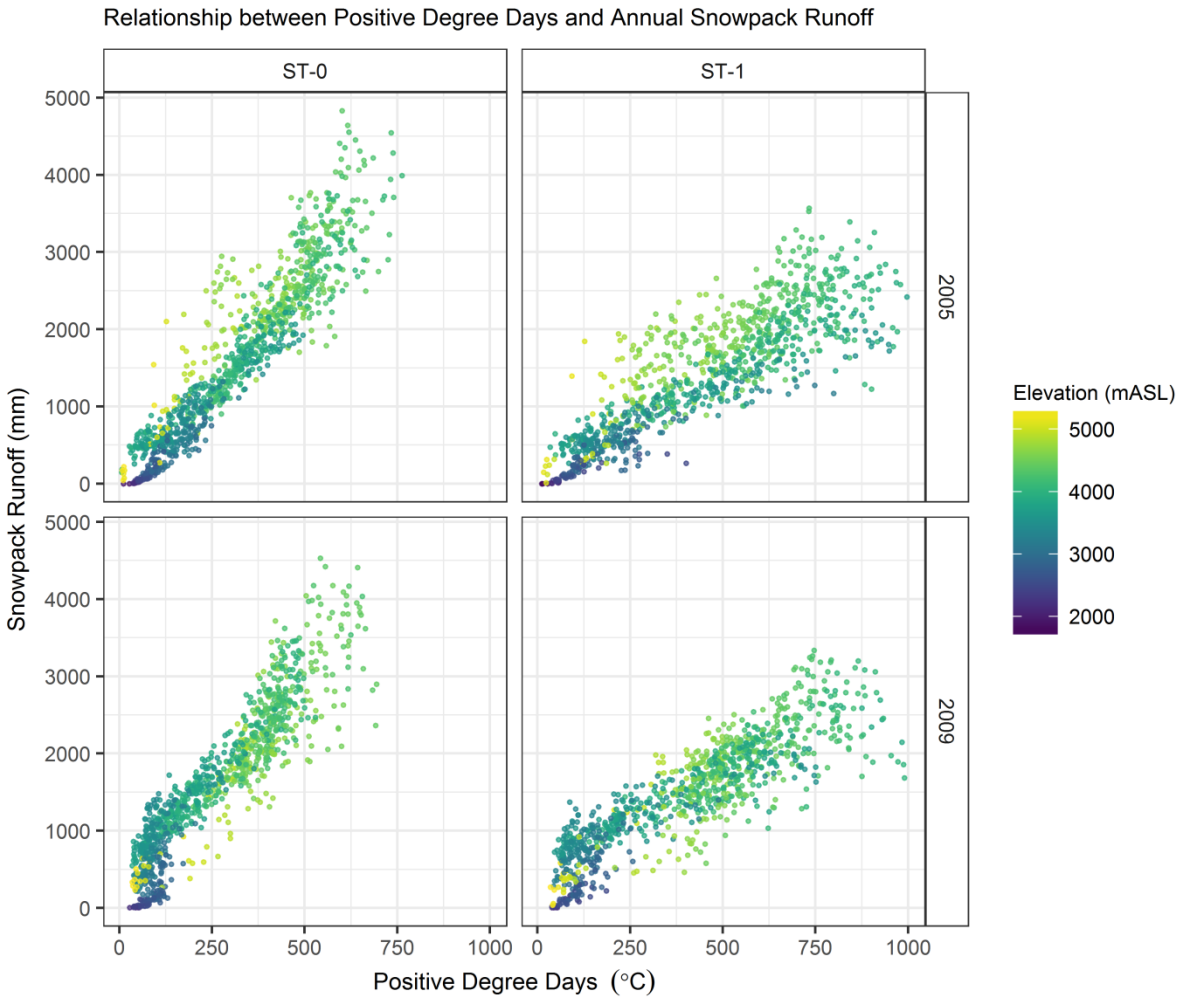


Figure 5.11 – Relationship between positive degree days (with snow cover) and annual snowmelt totals for two example years.

The results suggest that the relationship between cumulated positive air temperatures and snowpack runoff rates in FSM varies spatially (vertically) and between years. This is confirmed by Figure 5.11, which shows the relationship between cumulated positive degree days when snow is present, total annual snowpack runoff and elevation for each model cell in the Astore sub-basin. Over most of the elevation range, total annual snowpack runoff is close to peak snow water equivalent (SWE), after any sublimation losses in the period prior to complete ablation. Figure 5.11 shows that low elevations are subject to some non-linearity, increasing slowly at first then more rapidly. This might relate to the shallower, warmer snowpack at low elevations ripening relatively quickly, with rising net shortwave radiation (aided by decaying albedo) then rapidly inducing snowpack melt. At moderate elevations the relationship between positive degree days with snow cover and snowpack runoff is more linear, fitting better with the TI formulation. However, there is some suggestion that the highest elevations permitting melt show a different gradient in the relationship between

cumulative positive degree days (with snow cover) and snowpack runoff. Moreover, this high elevation gradient appears to differ between years. In 2005, the gradient is shallower than that at more moderate elevations, while in 2009 the gradient appears steeper than for moderate elevations. As the higher elevations undergoing melt in Figure 5.11 are still below the critical elevation for precipitation inversion (Section 5.2.1), this might suggest a particularly strong dependence of DDFs on inter-annual climate variability at high elevations, with implications for the high elevation water balance and ultimately glacier mass balance.

ETI Elevation Profiles – Threshold Temperature

The preceding analysis indicates that typical parameter values associated with the TI and ETI models can lead to varying levels of agreement with FSM in terms of catchment-scale snowpack runoff generation. Crucially, the consistency with FSM for any given parameter choice (held constant through the simulation) varies notably in time, on both intra-annual and inter-annual time scales. One additional parameter requiring consideration is the temperature threshold above which melt is permitted (Tt). To exemplify the influence of the Tt parameter, Figure 5.12 shows elevation profiles of ETI and FSM simulated snowpack runoff for two values of Tt, which are based on the ranges discussed in Ragettli et al. (2013, 2015). The baseline ETI configuration following the TOPKAPI-ETH setup is applied here, while the snowpack runoff shown in Figure 5.12 is from whole water years, defined as October to September. Thus, in the lower half of the elevation range there is essentially no difference between the models, as all snowfall is ablated over the course of a year. For ST-1 there may be some patches of persistent snow at relatively low elevations, perhaps on north-facing slopes.

The main difference between the ETI and FSM formulations occurs at higher elevations, where the ETI model experiences higher snowpack runoff rates. This suggests that the ETI formulation permits more rapid melt of the high elevation snowpack, getting into the relatively small fraction of the Astore catchment with perennial snow and ice. At this elevation, ablation in FSM is much lower and becomes dominated by sublimation. This could potentially be a manifestation of the earlier point, whereby the relationship between temperatures and release of snowpack runoff depends in part on snow depth, which is generally larger at high elevations as a function of simulated precipitation. Although the fraction of catchment area at high elevations is small, the high elevation water balance is important for the long-term evolution of the perennial snow and ice. Thus, while of lesser

significance for catchment-scale runoff total volumes over short time horizons, longer term trends in simulation of high elevation snowmelt will affect glacier mass balance and ultimately the degree of buffering of dry years by glacier melt, which can be significant even in Astore but especially in more glaciated areas (see below). These results suggest that parameters pertaining to melt onset delays in TOPKAPI-ETH may be really quite important, despite the fact that they did not appear in the sensitivity analysis results in Section 5.3.2.

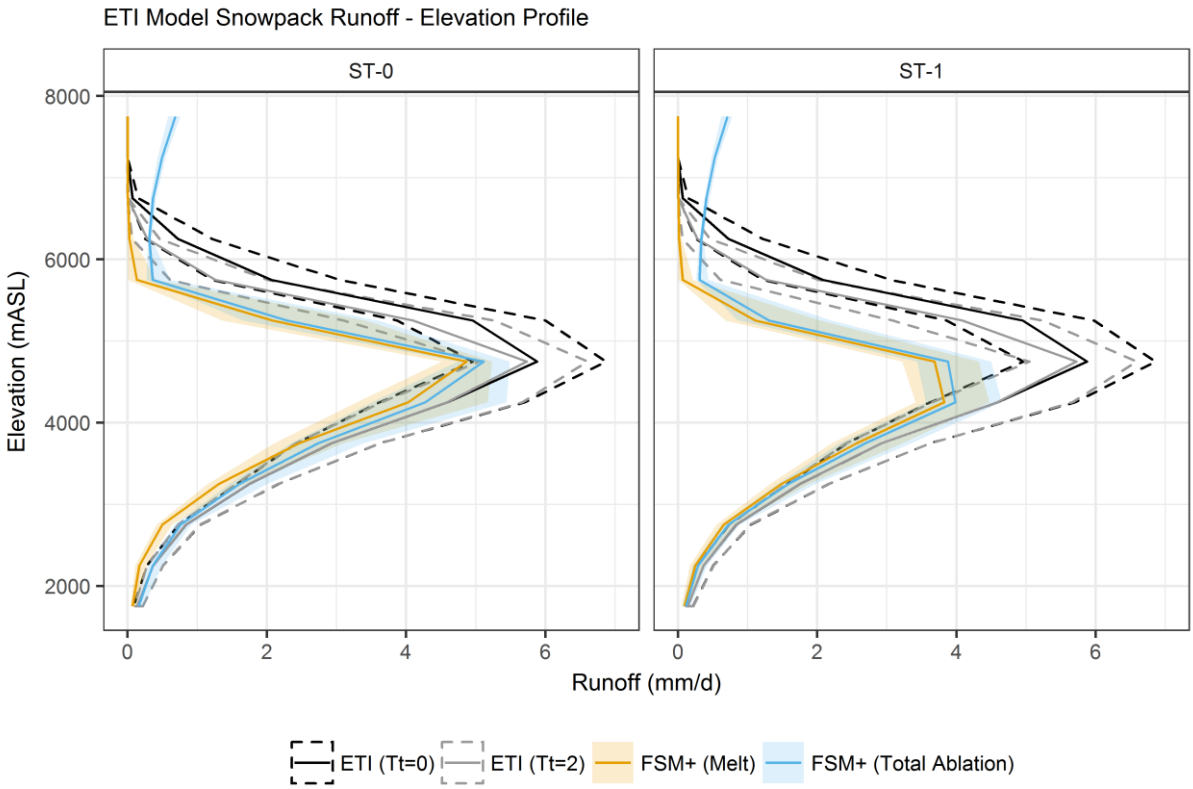


Figure 5.12 – Elevation profile of snowpack runoff (total over a year) in the baseline ETI model (i.e. baseline shortwave and temperature factors) but for different temperature thresholds compared with FSM. Both snowpack runoff released as liquid (i.e. melt) and total ablation are shown (i.e. additionally due to sublimation from the snowpack). Solid lines show means and ranges show inter-annual variability.

5.3.6 Simplified Energy Balance (SEB) and ETI Formulations

While the ETI melt model attempts to separate the temperature- and shortwave radiation-dependent terms of the surface energy balance equation to build on the simple TI approach (Pellicciotti et al. 2005), there does not appear to have been much investigation of whether the approach taken represents the best strategy. To explore this we compare the ETI model with a very similar formulation, namely the simplified energy balance (SEB) to ablation modelling

proposed by Oerlemans (2001) (equation 4.18 therein). This model was assessed by Gabbi et al. (2014) for glacier melt in the Alps, as well as MacDougall et al. (2011) for the St Elias mountains in the Canadian Yukon. In both cases, the SEB model appeared to perform roughly as well or indeed better than the ETI approach in transferability tests.

The SEB model following Oerlemans (2001) takes the following form:

$$Q_m = SW_n + c_0 + c_1 T_a \quad (4)$$

where Q_m is energy available for surface melt (Wm^{-2}), SW_n is net shortwave radiation (Wm^{-2}), c_0 is a parameter (Wm^{-2}), c_1 is a second parameter ($Wm^{-2}K^{-1}$) and T_a is air temperature (K). This can be rearranged simply as follows:

$$Q_m - SW_n = c_0 + c_1 T_a \quad (5)$$

If the ETI model is re-written to give melt in terms of energy, rather than water equivalent, it takes the following form:

$$Q_m = a\, SW_n + b\, T_a \quad (6)$$

where a is a dimensionless factor and b is a parameter ($Wm^{-2}K^{-1}$). With this formulation, dividing by SW_n suggests that the following relationship would take a linear form:

$$\frac{Q_m}{SW_n} = a + b\, \frac{T_a}{SW_n} \quad (7)$$

with each of the three terms being effectively dimensionless.

The linearity of the relationships proposed in equations 5 and 7 can be tested using outputs from FSM+, which computes melt energy based on solving the full surface energy balance. Figure 5.13 shows the relationships between the right- and left-hand sides of equation 5 for the SEB formulation and equation 7 for the ETI formulation, as calculated from FSM+ outputs. The FSM+ outputs are based on selecting a random sample of cells with snow melt at each time step, but where the snowpack has not completely melted (i.e. some snow mass remains at the end of the time step). The sample plotted in Figure 5.13 is for April to June, although similar patterns are found if the results are plotted for individual months separately

or for months later in the melt season. For each member of the sample, the melt energy, net shortwave radiation and air temperature from FSM+ were used to calculate the relevant terms of equations 5 and 7. The right- and left-hand sides of the equations are plotted on the vertical and horizontal axes, respectively.

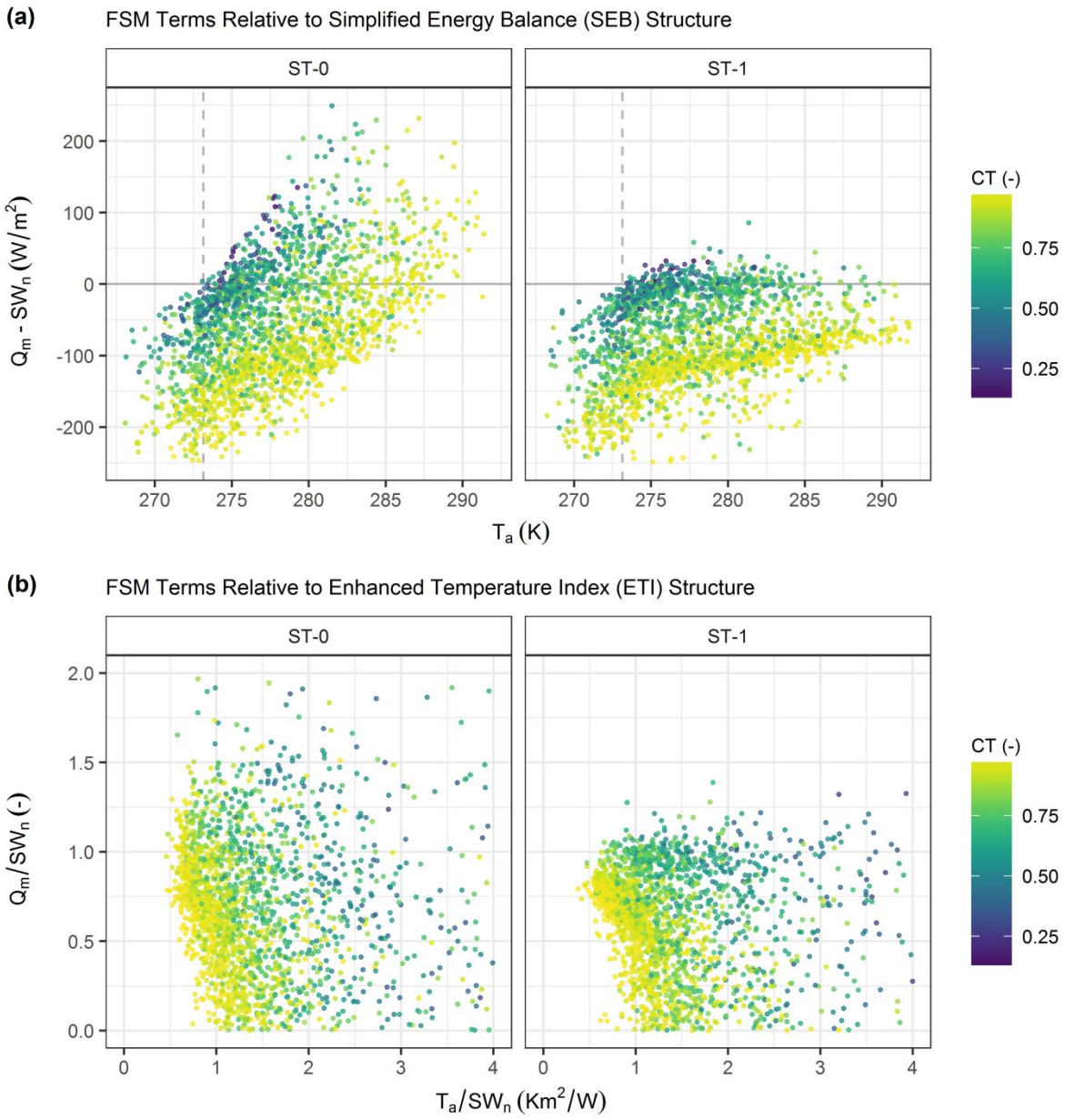


Figure 5.13 – Comparison of simplified energy balance (SEB) and ETI formulations in terms of implied relationships between temperature (T_a), net shortwave radiation (SW_n) and melt energy (Q_m). Each point is taken from one model cell selected randomly at each time step from the sample of cells where melt takes place but not all of the snow is ablated. The sample is for April to June. The comparisons for the SEB and ETI structures are given in (a) and (b) respectively, while the rationale for each is explained in the text. The dashed line in (a) marks 273.15 K.

For the SEB formulation, Figure 5.13a generally shows the linearity suggested by equation 5, but several complicating factors are evident. The most notable point is that there is additionally a strong dependence of the relationship between T_a and $Q_m - SW_n$ on cloud transmissivity (CT). Cloud transmissivity is defined here as the fraction of actual incoming shortwave radiation at the surface relative to that which would occur under clear-sky (i.e. cloud-free) conditions. In the ST-0 case, it appears that the colour gradient denoting cloud transmissivity is oriented approximately perpendicular to the overall positive relationship between T_a and $Q_m - SW_n$. In effect this suggests that the c_1 parameter is approximately constant, but that the c_0 parameter varies as a function of cloud transmissivity. The pattern in the ST-1 case in Figure 5.13a is similar in general, although notably the scatterplot appears somewhat truncated in the case where $Q_m - SW_n$ is close to zero. To some degree this gives the impression of possible non-linearity for cloudy conditions denoted by low cloud transmissivity, induced by the probably over-dampening of the sensible heat flux contribution to melting under stable conditions. In contrast, the relationship is strongly linear over most of the temperature range under relatively clear conditions. Some divergence from this occurs around freezing, where the gradient appears to steepen under all cloud conditions.

Oerlemans (2001) noted that c_1 could be fixed at around 10 $Wm^{-2}K^{-1}$. If c_1 is estimated from Figure 5.13a by calculating the gradient for, say, clear-sky (high cloud transmissivity) conditions, values of around 10 and 5 $Wm^{-2}K^{-1}$ are obtained for the ST-0 and ST-1 cases, respectively. These values are thus broadly in agreement with Oerlemans (2001), but highlight the notable effect of stability adjustment assumptions on the form of this relationship. Oerlemans suggested that c_0 could be a calibration parameter. In Oerleman's example for Nigardsbreen glacier in Norway, a value of -10 Wm^{-2} was used. Interestingly, this value appears roughly in agreement with the results in Figure 5.13a for cloudy conditions. Under clearer conditions, a much lower value of c_0 would be inferred from Figure 5.13a. In their application to long-term melt modelling for Rhonegletscher in Switzerland, Gabbi et al. found that a value of 15 $Wm^{-2}K^{-1}$ provided optimal results with respect to observations for c_1 over a multi-year calibration period, while a value of -75 Wm^{-2} was considered most appropriate for c_0. This latter value is within the range of approximately -200 to ~0 Wm^{-2} that might be inferred from Figure 5.13a, depending on cloud conditions, such that it may in effect be a compromise value obtained by not making c_0 explicitly dependent on cloud transmissivity. The climatic context of the UIB (lower latitude than the Alps, with higher clear-sky incoming shortwave radiation) may mean that the best compromise differs between locations.

Figure 5.13b shows that there are some similarities for the ETI formulation. Most notably, this includes an approximately linear relationship under clear-sky (high cloud transmissivity conditions). The major difference relative to the SEB case in Figure 5.13a is that the strength of the relationship breaks down very substantially as cloudiness increases. While the relationships in Figure 5.13a representing the SEB formulation remain comparatively tight as cloud transmissivity decreases, the relationships in Figure 5.13b representing the ETI formulation decay towards essentially showing very limited correlation between the constituent terms. This is consistent with the previous findings that the ETI formulation tends to work less effectively under cloudy conditions (Carenzo et al. 2009), which questions its applicability in particular climate settings to some extent. As cloud variation and trends are understood to be critical to ablation variability and climate change trajectories in the UIB (Forsythe et al. 2015), these results suggest that there may be important limits to the applicability of the ETI model in this context.

5.3.7 Glacier Ablation

Figure 5.14 shows daily snowpack runoff, glacier melt and total melt simulated by FSM+ along with observed total runoff for the heavily glaciated Hunza sub-basin. Total melt from FSM+ is the sum of snowpack runoff and glacier melt. Note that snowpack runoff refers to water discharged at the base of the snowpack, while glacier melt refers to surface melt, i.e. refreezing of glacier melt or glacier hydrology are not incorporated at this point. As per Section 5.3.4, rainfall runoff and evapotranspiration from model cells without snow or glacier present are not included here. Flow routing is also omitted here to focus on the runoff generation rather than routing processes.

Several points are evident from Figure 5.14. Firstly, there appears to be some strong similarities between the total melt series simulated by FSM+ and observed runoff, albeit with some complications. The degree of agreement can vary between and within years. The ST-0 case generally leads to higher total melt (through additional glacier melt), which leads to a series closer to observed runoff in some years (e.g. 2001) but not in other years, when peak melt rates may be notably overestimated. Quite plausibly, snowpack runoff leads glacier melt in the time series in all years, with the rate of rise and total magnitude of glacier melt quite substantially affected by the stability adjustment option selected. This is closely linked to the effects of stability adjustment on the rate of snowpack melt-out. With ST-1, snow overlying glacier cells may persistent for longer, thus delaying the onset of significant glacier melt. The

timing of intra-season fluctuations in snow/glacier melt generally match well with observed runoff. Overall then it appears that FSM+ provides potential for simulating catchment-scale glacier melt reasonably.

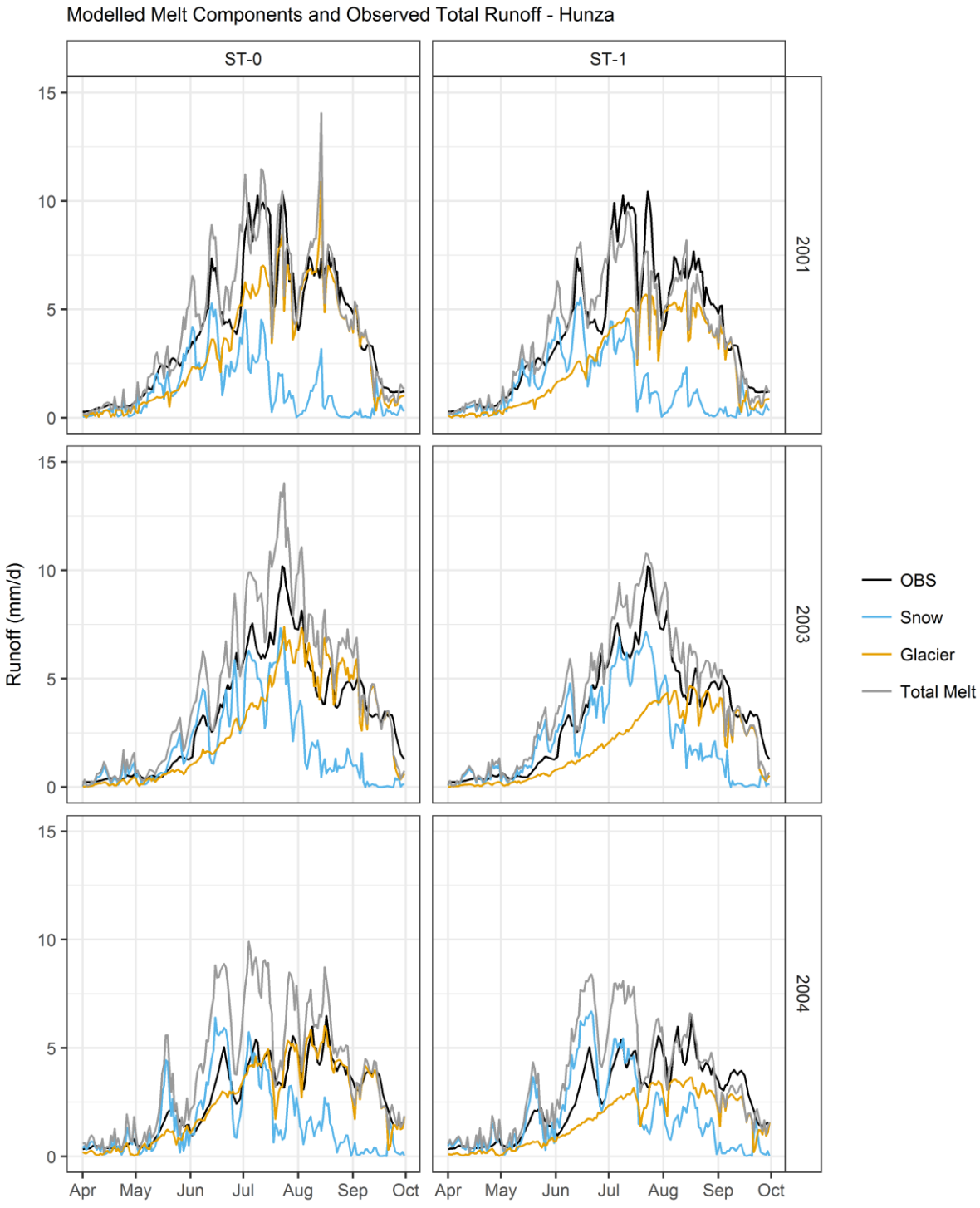


Figure 5.14 – Comparison of daily snowpack runoff and glacier melt time series simulated by FSM+ with observed runoff for the Hunza basin for selected contrasting years. Snowpack runoff (blue), glacier melt (orange) and total melt (grey) simulated by FSM+ are not routed (and runoff and evapotranspiration from cells without snow or glacier present are not considered).

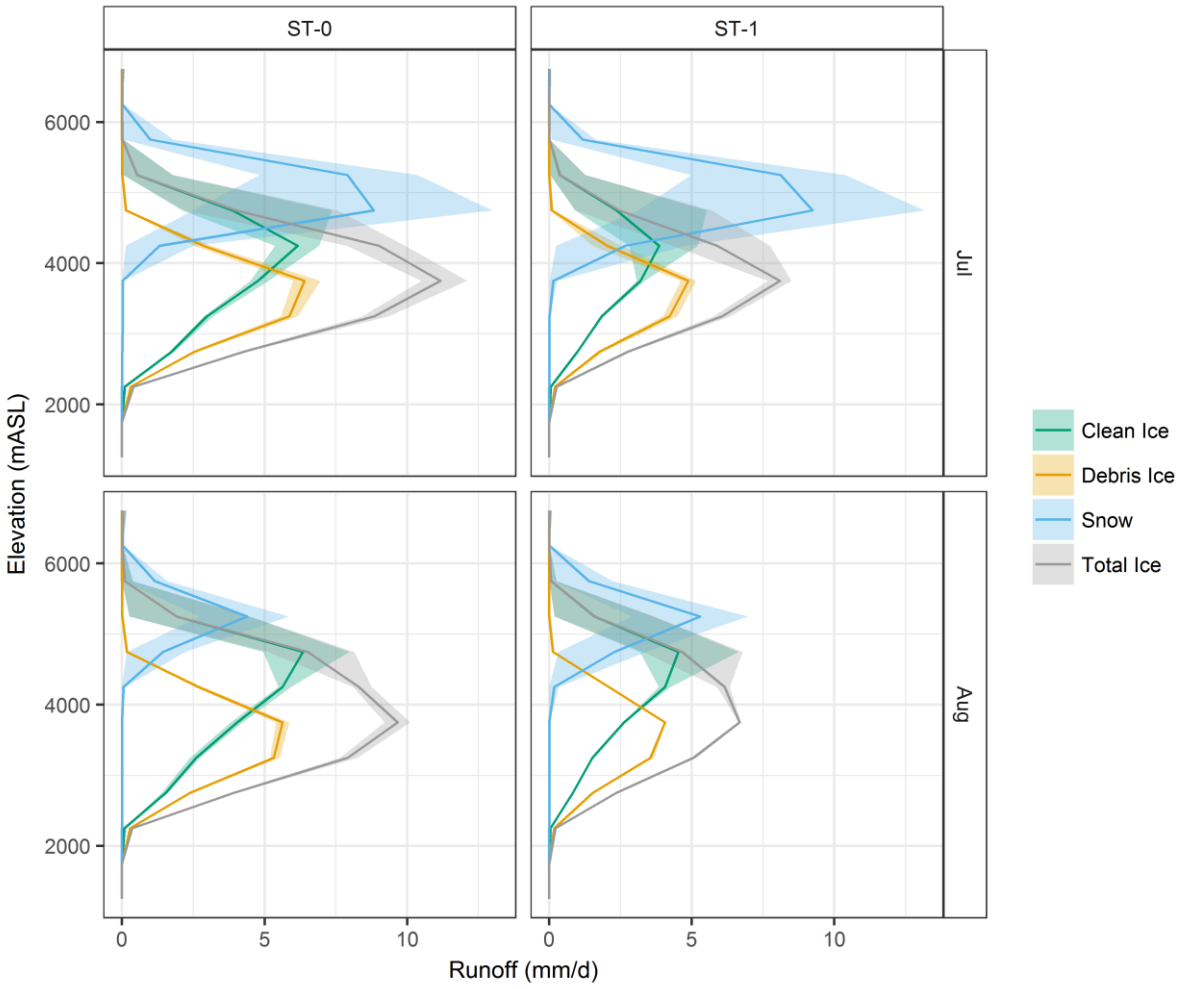


Figure 5.15 – Elevation profiles of melt components for the Hunza sub-basin. Lines show means across years, while shading denotes the range of inter-annual variability.

Figure 5.15 shows elevation profiles of snowpack runoff and glacier melt for the Hunza sub-basin for July and August, the two months typically having highest glacial runoff. Several similarities are apparent in both months and between stability adjustment options. Firstly, the peaks in sub-debris melt, bare ice melt and snowpack runoff are at successively higher elevations in a given month. The bare ice melt and snowpack runoff components peak at higher elevations in August than July. Secondly, the fraction of total glacier melt occurring at lower (but partly glaciated) elevations through sub-debris melt is notable. This is likely related to the general pattern of debris thickness increase towards glacier termini, where debris is present, such that fairly substantial melt rates can still occur higher up the glacier tongues and ablation areas where debris cover is thinner. Thirdly, inter-annual variability in glacier melt below the total ice melt peak is fairly limited for both the bare ice and sub-debris cases, especially in August. In the case of debris cover, this suggests that debris thickness is likely to be a more important control on melt rates than climate variability where debris is

present. The minimum temperature isotherm averages just under 5000 mASL in July and August, so it is reasonable that substantial energy is available for ablation in these months.

Above 4000 mASL bare ice melt comes to dominate total ice melt, which is accompanied by a notable increase in inter-annual variability. This increase in variability coincides with the rise in relatively high elevation snowmelt. This has the largest variability of any of the melt components, largely as a function of precipitation variability in the preceding winter and spring. The results suggest that the persistence and rate of snowmelt strongly affect the simulated partitioning between snow and glacier melt in the higher part of the ablation zone. In July snowmelt tends to substantially exceed ice melt from around ~5000 mASL upwards, whereas by August the difference decreases owing to the depletion of the snowpack.

While local data for full evaluation of the ablation from clean and debris-covered glacier ice are not available, MODIS LST provides one means of partially investigating how well FSM+ simulates the surface energy balance. Details of MODIS LST validation are given in Section 5.2.3. Figure 5.16 shows elevation profiles of day and night surface temperatures (mean and inter-annual range – based on MODIS overpass times) for the main months of high glacier melt in the Hunza sub-basin. The temperatures are averaged over all glaciated cells or pixels within the elevation band, i.e. both clean and debris-covered.

For day-time, this suggests that surface temperatures are well simulated in the upper reaches of glaciers under both stability options. MODIS LST then tends to increase above freezing at a slightly higher elevation than in FSM+, but the latter does indeed capture the transition to positive temperatures reflecting debris presence. This may suggest that debris cover extends higher in some places than inferred from Khan et al.'s (2015) mapping. However, at elevations below around 4000 mASL, surface temperatures tend to diverge from MODIS LST, which continues to increase. Possible explanations for this include uncertainties in debris thickness or debris bulk properties. Interestingly, Collier et al. (2015) show similar results in their equivalent elevation profile (their Figure 3a), whereby divergence from MODIS is highest between 5000 and 6000 mASL and then again at the lowest glaciated elevations.

At night, Figure 5.16 becomes dominated by the discrepancy arising from stability option choice at higher elevations associated with clean ice conditions. With ST-0, reasonable agreement between FSM+ and MODIS is attained, but again FSM+ struggles to match the positive temperatures at low elevations suggested by MODIS.

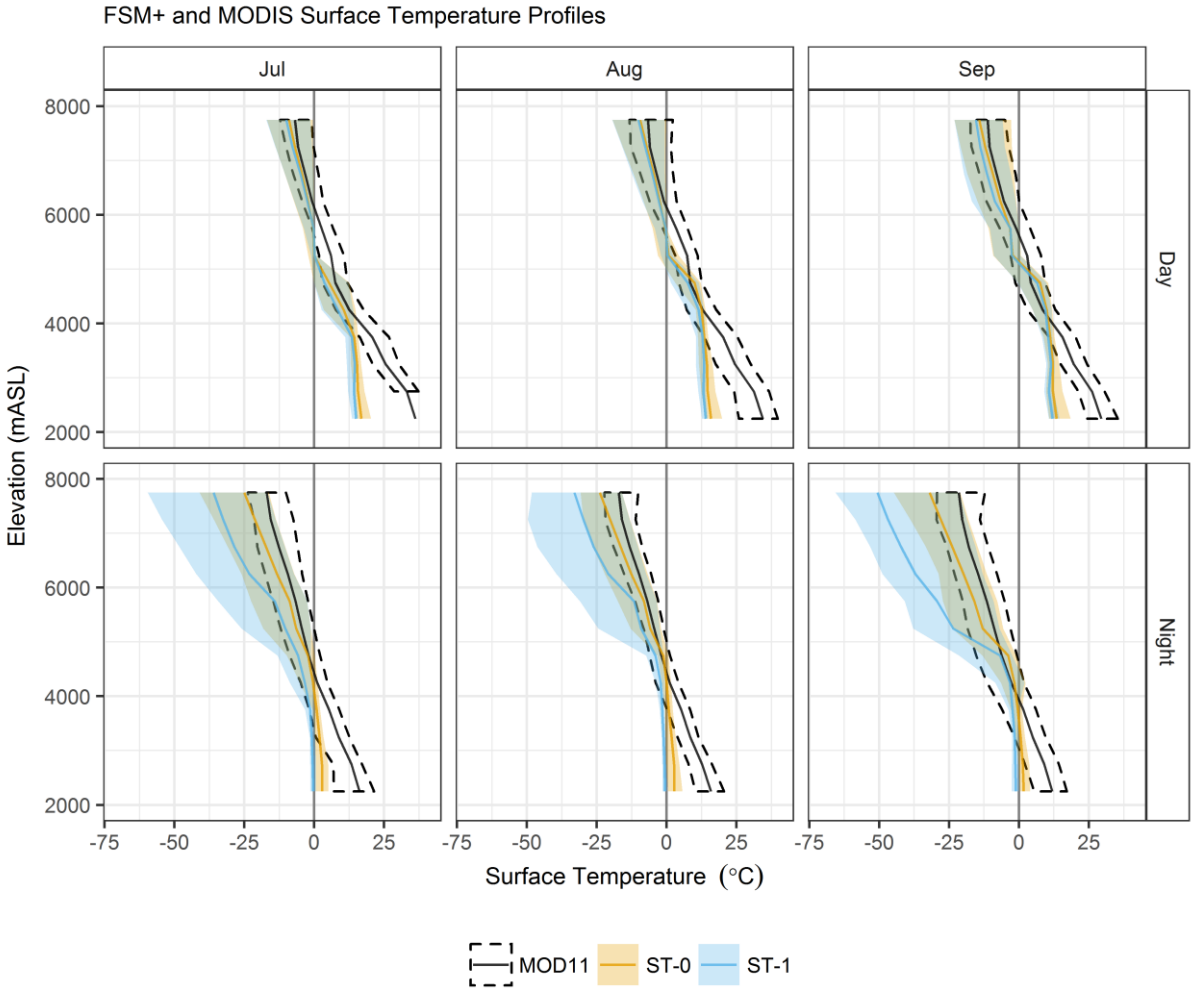


Figure 5.16 – Comparison of glacier surface temperatures in FSM+ and MODIS.

5.3.8 Glacier Clean Ice Ablation Comparison

In the glacier melt model comparisons reported below, ice ablation using these empirical approaches (i.e. both clean ice and sub-debris melt) begins in a cell following ablation of the snowpack, which is simulated using the full surface energy balance approach of FSM+. This is to help isolate the glacier melt responses of the empirical parameterisations, without the confounding influences of different rates of snowpack ablation from alternative snow melt model options.

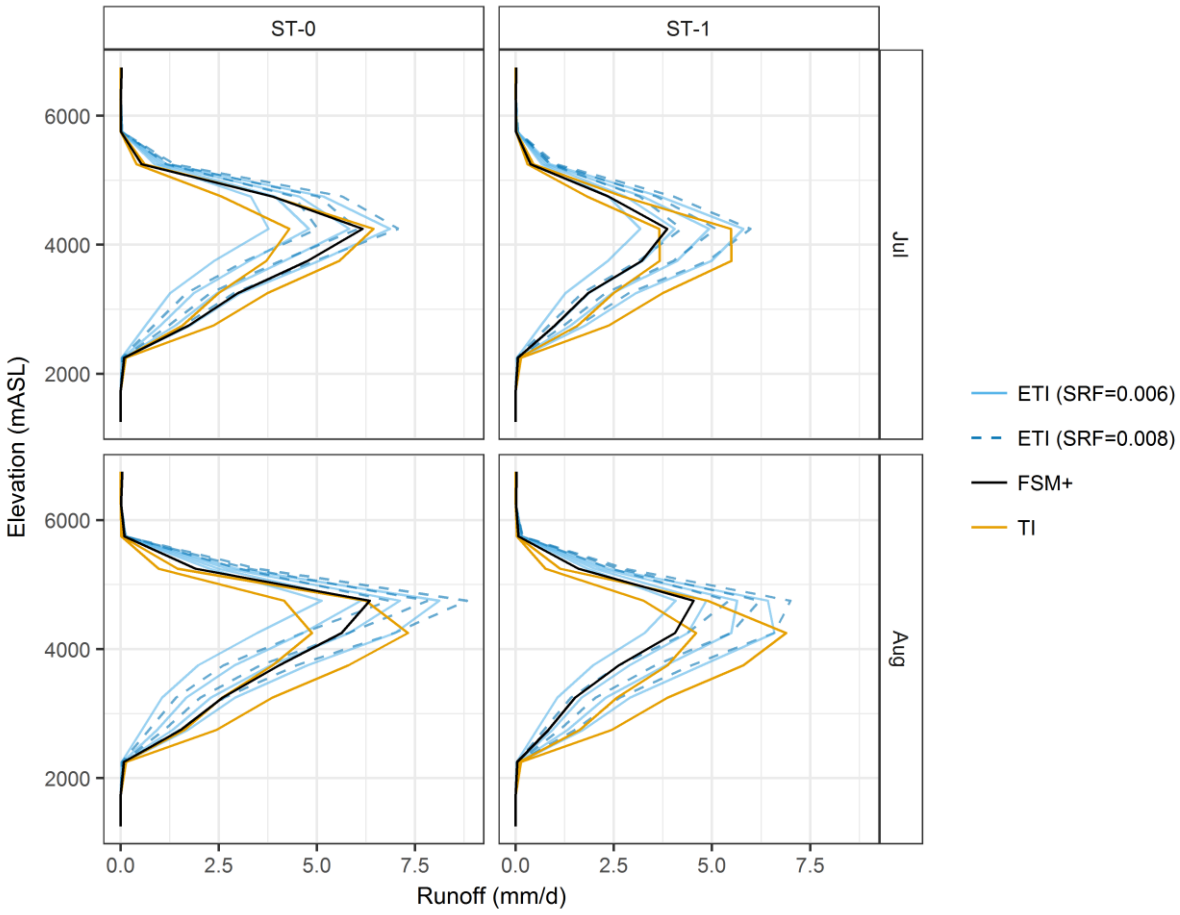


Figure 5.17 – Elevation profiles of glacier clean ice melt for FSM+ and TI (orange) and ETI (blue) empirical methods. For the ETI model, the cases of two separate shortwave radiation factors (SRF) are denoted by solid and dashed lines. For the lower SRF (0.006 m^2 mm W^{-1} hr^{-1}), profiles for temperature factor values of 0.0, 0.04, 0.08 and 0.12 mm hr^{-1} $°C^{-1}$ are shown (each leading to successively higher melt rates), whereas profiles of temperature factor values of 0.0, 0.04 and 0.08 mm hr^{-1} $°C^{-1}$ are shown for the higher SRF case (0.008 m^2 mm W^{-1} hr^{-1}) (again each leading to successively higher melt rates). The TI models shown are for degree-day factors of 4 and 6 mm d^{-1} $°C^{-1}$.

Figure 5.17 compares bare ice glacier melt calculated by FSM+ with that from TI and ETI models. While a range of parameters were tested for the TI and ETI models, only those close to the FSM+ outputs are shown to boost clarity. Focusing first on the shape of the profiles, it seems that the ETI model tends to agree more closely with the shape of the FSM+ profiles. This is because melt rates in the TI profile tend to increase relatively rapidly at lower elevations, compared with FSM+ and ETI profiles. Thus, there is some suggestion that the ETI functional form could be more appropriate. However, there are significant complications regarding parameter values. For the TI model, the DDF leading to most agreement with FSM+ varies with elevation, with higher DDFs being implied at higher elevations. For the

ETI model, the close overlap of the dashed and solid lines on Figure 5.17 demonstrates the high possibility of parameter compensation. Specifically, increasing the shortwave radiation factor (SRF) by 0.002 m^2 mm W^{-1} hr^{-1} in Figure 5.17, and at the same time decreasing the temperature factor (TF) by 0.04 mm hr^{-1} $°C^{-1}$ (down to a minimum of zero), leads to essentially very similar elevation profiles of bare ice melt. The ST-0 case generally implies higher melt, but interestingly cuts across some of the profiles for different parameter combinations, suggesting some variation of implied parameters with elevation.

5.3.9 Debris-Covered Ice Ablation Comparison

To examine the simulation of sub-debris glacier melt in FSM+, we compare daily time series with several empirical approaches in Figure 5.18. The first point to note is that there is notable difference between ST-0 and ST-1 series in FSM+, with the latter exceeding the former quite significantly. It is possible that a substantial portion of this difference relates to effects of stability adjustment on the diurnal cycle of melting. The degree of debris surface cooling at night-time might be somewhat limited in ST-0 (Figure 5.16), allowing higher debris temperatures to be maintained and so greater sub-debris melt. It is also clear that there is a fairly high degree of agreement between the FSM+ (ST-1) and the TId and ETId formulations using baseline parameters. ETId appears to show slightly less melt than the other two models, but the difference is relatively small compared with the overall magnitude of runoff at the catchment scale and the various sources of error involved.

Interestingly, the two simulations with the DETI parameterisation approximately bound the other approaches. Transferring parameters from Carenzo et al.'s (2016) model development in the Alps leads to lower sub-debris glacier melt relative to transferring parameters from Ragettli et al.'s (2015) application in the Langtang valley (Nepal), which matches closely with FSM+ with no stability adjustment. In both of these studies, parameters were estimated by calibrating to outputs from an energy balance model accounting for heat conduction through supraglacial debris (Reid and Brock 2010), which led to good overall model performance in their test cases. Ragettli et al.'s application included debris thicknesses in excess of 2 m, while Carenzo et al.'s application concerned debris up to 0.5 m thick (with most of their measurements and experiments for debris layers of up to 0.2 m).

One possible reason for the large differences relative to the FSM+ cases and the TId/ETId empirical approaches might lie in the structure of the DETI formulation. In effect the DETI method has 7 parameters compared with the standard ETI's 3 parameters (including a

temperature threshold for melt). Several of the DETI parameters are calculated as nonlinear functions of debris thickness, such that its sensitivity to small changes in parameter values may be large. The structure is clearly flexible enough to mimic an energy balance model following calibration, but without calibration its transferability to different contexts could be quite limited.

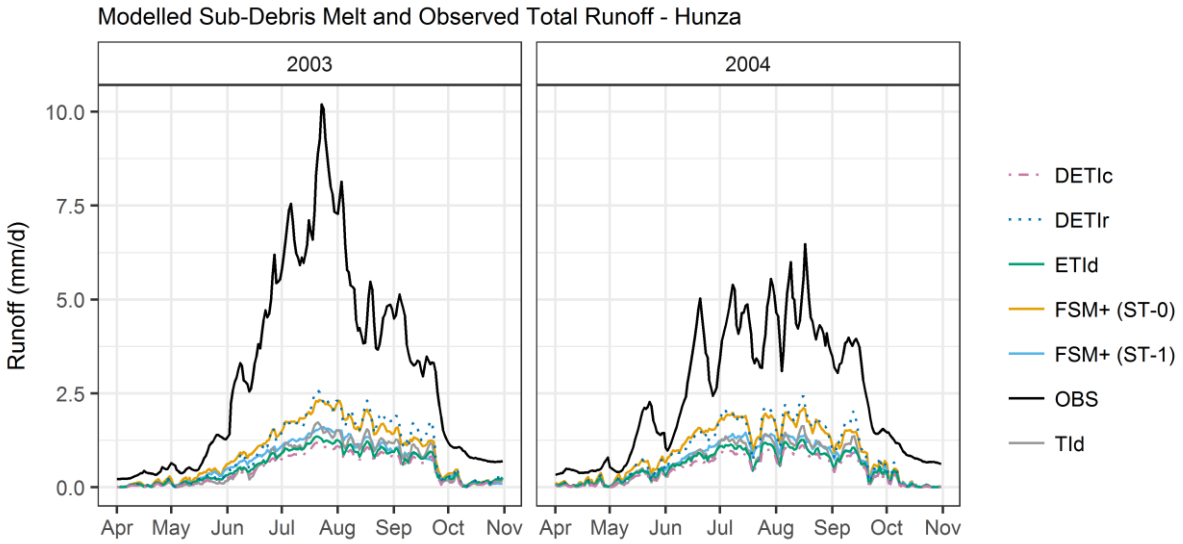


Figure 5.18 – Comparison of FSM+ and empirical approaches for modelling melt under debris cover in the Hunza sub-basin. Observed total runoff is shown for reference in black. The empirical approaches are: DETIc – DETI method accounting for debris thickness using parameters from Carenzo et al. (2016); DETIr – DETI method using parameters from Ragettli et al. (2015); ETId – ETI method with binary debris cover and modified ETI parameters following Ragettli et al. (2013); TId – TI method with binary debris cover using parameter from Lutz et al. (2016a).

5.3.10 Glacier Ablation Temperature Sensitivity

Figure 5.19 shows the results of a simple sensitivity test for the TI, ETI and FSM+ clean ice glacier melt formulations, in which the input air temperature time series was perturbed by constant increments. The TI and ETI results shown are based on parameter values (or combinations in the ETI case) where catchment-wide mean annual clean ice melt was within 10% of the respective FSM+ model using the unperturbed input air temperatures. As none of the regular interval DDFs tested for the TI model were within 10% for the ST-1 case, this range was expanded to 20% in Figure 5.19b for the TI model only.

Intuitively, Figure 5.19 suggests that the TI formulation exhibits the highest sensitivity to temperature perturbations. Unlike the other models, the TI approach depends only on

temperature. As such, rises in temperature lead to comparatively steep increases in melt where ablation is not mass-constrained, such as in the heavily glaciated Hunza sub-basin under typical conditions. In Figure 5.19b, the shallower of the two TI lines corresponds in fact with catchment-wide mean annual clean ice melt that is around 20% lower than in FSM+ in the unperturbed model run. Yet, its temperature sensitivity remains much higher than FSM+ with ST-1. With a perturbation of +3°C, this formulation results in approximately 100 mm/a of clean ice melt, compared with around 50 mm/a in the equivalent FSM+ scenario.

For both ST-0 and ST-1, four (different) parameter combinations provided mean annual clean ice melt within 10% of the FSM+ simulated totals. In both cases, three of these parameter combinations remain relatively closely grouped together. These groups show the lowest temperature sensitivity of the empirical models and parameter values/combinations plotted. However, these groups also show notable differences from the FSM+ sensitivities, with the closely grouped ETI models being leading to clean ice melt around 25% and 50% higher than FSM+ in the ST-0 and ST-1 cases. In both the ST-0 and ST-1 cases, one of the ETI parameter combinations exhibits considerably more temperature sensitivity than the other three, as reflected by the steeper line in Figure 5.19. Clean ice melt is around 65% and 100% higher in the ETI model with these parameter combinations than FSM+ in the ST-0 and ST-1 cases, respectively. This highlights the importance of applying conceptual and process-based understanding to the ETI model parameters where insufficient data are available to truly test the ETI model for application at the catchment scale. To some extent it appears to be the modeller's choice how much temperature sensitivity is assigned to the ETI model.

Importantly, the differences between the ETI and FSM+ models appear to increase with magnitude of the temperature perturbation. This is also the case for the TI model. Even in the case of no stability adjustment in FSM+, offering the potential for substantial sensible heat flux to contribute to melt as the summer progresses, clean ice melt is more sensitive to temperature even in the more conservative ETI parameter combinations. It seems reasonable to suggest then that there is a strong possibility that the ETI formulation would lead to more rapid changes in glacier mass balance than the energy balance approach, at least in realistic (simplified) applications of climate change impact assessment. While differences in clean ice melt over a relatively short timeframe appear substantial within themselves, the impact could grow over time as the cumulative effect of higher/lower ablation plays out on glacier mass balance. It might be expected that river flows during dry years may be particularly affected by this. As glacier volume declines then so too does the buffering effect of glacier melt on inter-

annual climate variability, such that the differences between model formulations could ultimately become very important during times of water stress. This is discussed further in Chapter 7.

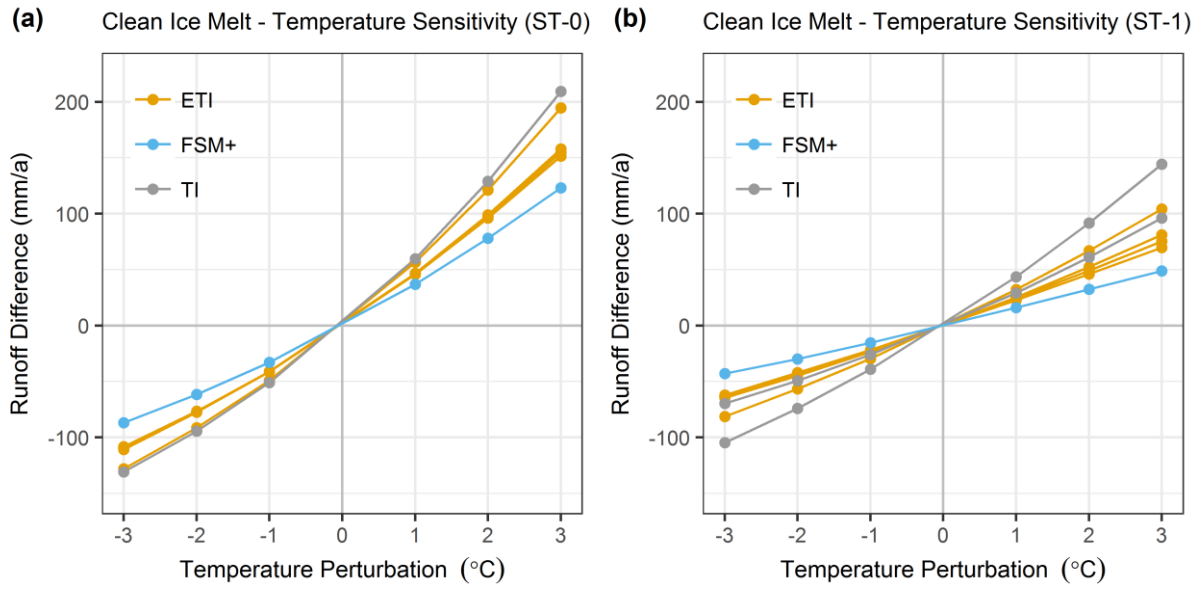


Figure 5.19 – Sensitivity of annual clean ice glacier melt in the Hunza sub-basin to simple temperature perturbations for TI, ETI and FSM+ models. The TI and ETI results use FSM+ to simulate the overlying snowpack, so total snowpack ablation occurs at the same time in the comparisons (but at different times in the ST-0 (a) and ST-1 (b) cases). The TI and ETI comparisons shown are for those parameter combinations producing catchment-wide mean annual clean ice melt within 10% of the respective FSM+ model with unperturbed temperatures. The exception is TI in (b), where the results correspond to +/- 20% of FSM+ in the unperturbed case, as the regular interval DDFs tested did not give any simulations within +/- 10%.

5.4 Conclusions

This chapter demonstrates that TOPKAPI-ETH configured following Ragettli et al. (2013) provides generally good performance relative to available observations. The parameter set is perhaps not quite optimal, but it seems to be a reasonable choice overall. This provides further confirmation of the value of the HAR for climate inputs when the full hydrological cycle is considered. While the benchmark TOPKAPI-ETH setup is reasonable, the chapter does also confirm that the ETI method used to compute ablation is subject to substantial parameter equifinality. There may be a case for reducing the parameter used to factor net shortwave radiation if the ETI model is to more closely mimic an energy balance approach. Otherwise

the potential ablation may be excessive, but overall we suggest that the benchmark parameter choices are reasonable and suitable for hydrological projections.

The results also show that there are some notable differences between TI and ETI formulations compared with the FSM model for snowpack runoff modelling. Substantial variation of parameters in both space and time would be required to mimic the response of FSM. This is particularly the case for simple models omitting processes of liquid water retention, drainage and melt/refreezing cycles, the implications of which may be critical for the long-term evolution of high elevation water balances. Temporally, this variation follows a systematic seasonal pattern in general, but there is substantial inter-annual variation. Longer time series may be required to gain a better characterisation of the interactions between climate input fields and the modulating effects of snowpack processes. Yet, the results suggest that notable complexity is lost in both the TI and ETI approaches relative to the energy balance model at the heart of FSM. This complexity may pertain to the multiple dimensions of the climate inputs and the various feedbacks possible within the snowpack. How significant these omissions in simpler models are for hydrological projections is unclear.

While the ETI model performs well based on available observations of river flows and SCA, its weaker performance under cloudy conditions may be a structural limitation. Specifically, the linear relationship between its inputs (temperature and net shortwave radiation) and output (melt) implied by its formulation breaks down with increasing cloud cover. In contrast, owing to a small difference in form, the simplified energy balance (SEB) method (Oerlemans 2001) retains a more stable relationship between the inputs and output under all cloud conditions. Albedo parameterisations aside, both the ETI and SEB formulations have two calibration parameters and the same input requirements. Thus the SEB approach could in theory be more robust. However, one of the parameters at least shows a dependence on cloud cover, which would need to be accounted for, perhaps with a simple parameterisation, unless an “average” value for all cloud conditions were deemed acceptable.

For glaciers, the ETI approach shows fairly good functional agreement with energy balance models for clean and debris-covered ice ablation. The TI model may exhibit some more differences in behaviour. Potential differences in functional form become particularly critical when considering the temperature sensitivity of different models. The TI model exhibits a much stronger sensitivity to temperature perturbations than the energy balance formulations. The magnitude of this difference is large and significant for water resources applications.

Some parameter combinations for the ETI model show much more agreement with the sensitivity of the energy balance approach. Although the simplicity of the sensitivity tests should be acknowledged, ETI-based models do thus appear to provide an advance over TI models when the former are appropriately parameterised.

Chapter 6

Controls on Temperature Patterns and Projections in CMIP5

6.1 Introduction

The substantial uncertainty in climate and hydrological projections for the Himalayan region poses huge challenges for water resources planning (Mishra 2015). One key contributor to this uncertainty is the large spread in GCM projections (e.g. Ragettli et al. 2013; Lutz et al. 2014, 2016a,b; Palazzi et al. 2015). In order to understand and potentially constrain this spread, it is first critical to evaluate how well GCMs simulate key features of the regional climate, and whether there are connections between model skill, process representations and projections. Previous studies have evaluated a number of pertinent aspects of GCM behaviour over the Himalaya and Tibetan Plateau. These include: precipitation and temperature climatology (e.g. Su et al. 2013; Palazzi et al. 2015; Hasson 2016a; Hasson et al. 2016); land- and near-surface processes, feedbacks and elevation-dependent warming (e.g. Rangwala et al. 2013; Ghatak et al. 2014; Rangwala et al. 2016; Palazzi et al. 2017; Guo et al. 2018); monsoon dynamics (e.g. Sperber et al. 2013; Levine et al. 2013; Saha et al. 2014; Wang et al. 2017); and other relevant facets of global and regional atmospheric circulation (e.g. Lee et al. 2014; Xu et al. 2017).

One climate feature that has not been investigated in GCMs is the Karakoram Vortex (KV) (Forsythe et al. 2017; Li et al. 2018). As a new and emerging research area, the KV shows considerable promise for understanding the drivers of contrasting climate change signals across the Himalaya and Tibetan plateau. A fuller description is provided in Section 6.2, but in brief the KV describes anomalous circulation patterns centred on or near the Karakoram throughout the year. Anticyclonic and cyclonic KV states are closely linked to anomalies in the position and intensity of the subtropical westerly jet (SWJ) in all seasons, as well as the South Asian monsoon (SASM) in summer. The strong, spatially extensive and homogenous temperature response to KV circulation anomalies in winter contracts substantially in summer, becoming focused on a more restricted zone around the Karakoram. In conjunction with trends in the SWJ, SASM and KV states, Forsythe et al. (2017) argued that this contraction during the peak ablation season provides a significant contribution to the unique summer cooling (Fowler and Archer 2006; Forsythe et al. 2012b), neutral glacier mass balance (e.g. Hewitt 2005; Zhou et al. 2017; Bolch et al. 2017; Brun et al. 2017), and stable or

declining river flow trends (Sharif et al. 2013) observed in the Karakoram and Upper Indus Basin (UIB).

Although the KV is a recently introduced concept, its role as a major control on regional temperature variability, particularly for the western Tibetan Plateau and adjoining mountain ranges (Li et al. 2018), is increasingly supported by other, independent studies. For the summer monsoon season, Mölg et al. (2017) demonstrated that southward shifts in mid-latitude westerly activity and the SWJ are associated with a dipole in temperature response, with cooling (warming) in the west (east) of the Tibetan Plateau. Norris et al. (2018) additionally showed that contrasting cyclonic (anticyclonic) trends in the western (central/eastern) parts of the Himalayan arc and Tibetan Plateau help to partly explain regionally differentiated temperature and glaciological trends. These findings are largely consistent with the circulation anomalies, temperature responses and trends associated with the KV (Forsythe et al. 2017; Li et al. 2018). Therefore, we contend that the KV represents a valuable means of investigating and quantifying the year-round impacts of SWJ variability and summer SWJ-SASM interplay on regional near-surface climate.

While KV representations and projections in GCMs have not yet been studied, a related summer circumglobal teleconnection pattern (CGT) (Ding and Wang 2005) has been assessed in the Coupled Model Inter-Comparison Project phase 5 (CMIP5) GCMs (Taylor et al. 2012). The CGT is associated with a global wave train closely guided by the SWJ, and the KV is broadly consistent with its west-central Asian action centre. Lee et al. (2014) showed that the CGT is reasonably well reproduced by some GCMs. From analysing Representative Concentration Pathway (RCP) scenario outputs, Lee et al. found that the CGT is projected to weaken under RCP4.5, while the relationships between the CGT and the SASM, as well as the SASM and the El Niño Southern Oscillation (ENSO), are also projected to decrease in strength. However, this analysis is restricted to summer, and it does not focus on GCM performance and projections in detail for the regional circulation and near-surface climate responses within the domain of the Himalaya and Tibetan Plateau influenced by the KV.

As such, we evaluate three aspects of KV representation using an ensemble of GCMs from CMIP5. Firstly, we examine how well the basic indices and structure of the KV are represented in the CMIP5 ensemble in different seasons. Secondly, we focus on the Karakoram and UIB to consider how well the historical connections between regional circulation and near-surface air temperature variability and trends are captured by the GCMs.

This is a key test of whether the models reproduce the patterns of spatial variation in temperature response that underpin the divergent climatic, hydrological and glaciological trends across the Himalayan arc. Finally, we consider the connections between KV representation and ensemble spread in temperature projections for the Karakoram and UIB, focusing on global temperature stabilisation scenarios. This leads us to consider other critical influences on uncertainty in projected changes, as well as the degree to which they may be constrained by observations.

6.2 Karakoram Vortex

In this section, we review some basic features of the regional atmospheric circulation and explain the KV, thus providing the context for evaluation of the CMIP5 models. The explanations are based on the detailed accounts in Forsythe et al. (2017) and Li et al. (2018) Figure 6.1a shows the climatological horizontal winds at 200 hPa for winter (December, January, February - DJF) and summer (June, July, August - JJA). The band of high zonal wind velocity denoting the SWJ is especially clear in winter, when the SWJ is located to the south of the Karakoram. Between winter and summer, the SWJ reduces in speed and migrates north of the Karakoram and much of the Tibetan Plateau. The line of minimum zonal wind speed in summer is coincident with the ridge line of the South Asian High. Lower level south-easterly flows from the Bay of Bengal and south-westerly flows from the Arabian Sea in summer are associated with the SASM.

Figure 6.1a also shows the northern and southern areas used to define the Karakoram Zonal Shear (KZS) index. To calculate the KZS, zonal wind speeds are first spatially averaged over the northern and southern areas separately. The KZS is then defined as the resulting northern area value minus the southern area value. Calculated in all seasons, the KZS quantifies the position and intensity of the SWJ upstream of, and around, the Karakoram. The KZS is negative in winter when the SWJ is at its southernmost position and zonal wind speeds are higher in the southern area. KZS then increases to a maximum in summer, becoming positive and reflecting the northward migration of the SWJ. The majority of the amplitude in the KZS annual cycle derives from seasonal variation in the southern area.

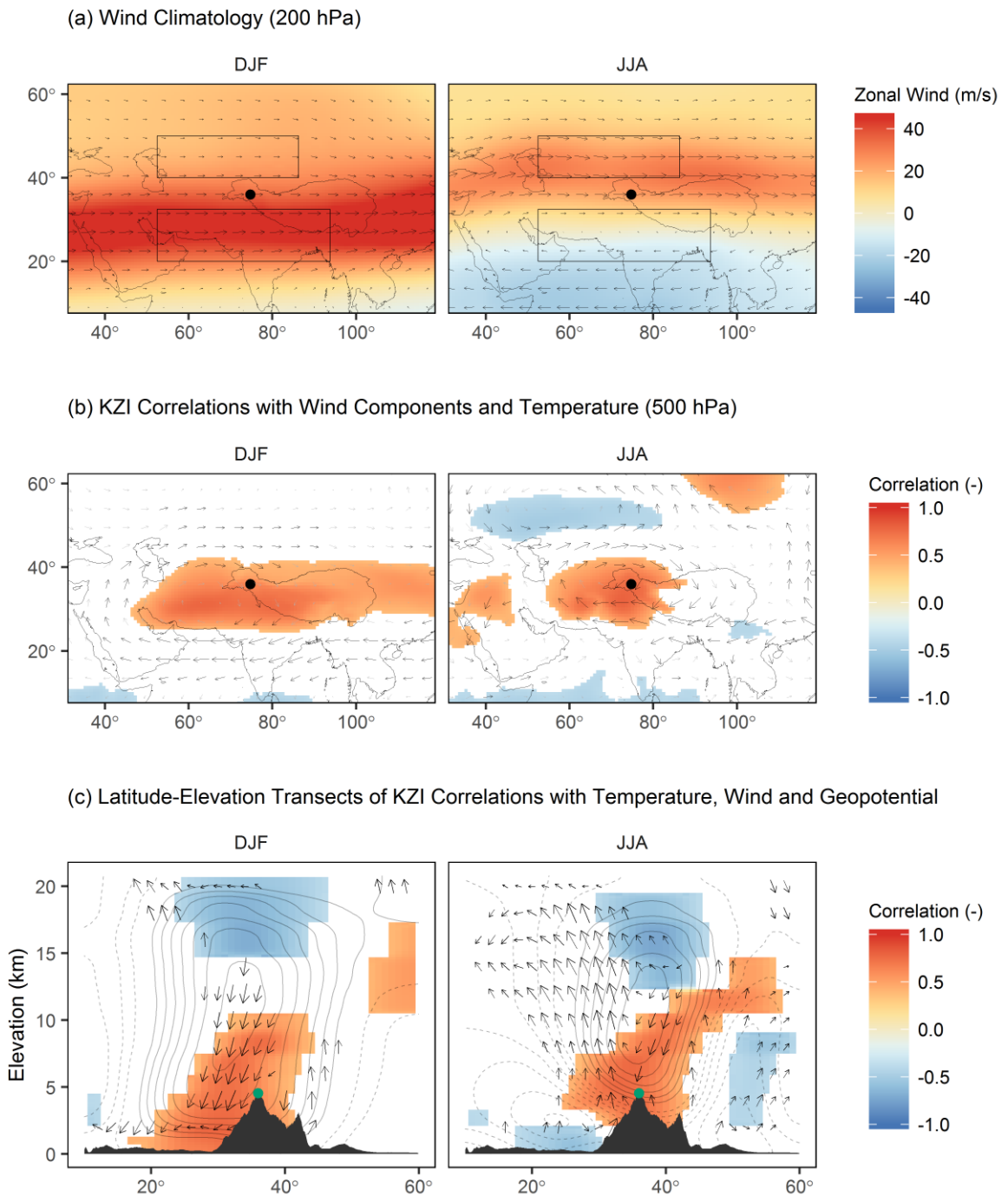


Figure 6.1 – Summary of the Karakoram Vortex (KV) based on ERA-Interim (ERAI) 1979-2005 for winter (DJF) and summer (JJA). (a) Wind climatology at 200 hPa, with shading showing zonal wind speed. Boxes used to define Karakoram Zonal Shear (KZS) index are also shown (see Section 6.2). (b) Correlation of KZI with temperature (shading) and horizontal wind components (arrows). (c) Latitude-elevation transects (integrated over 70-80°E) of KZI correlations with temperature (shading), meridional and vertical wind (arrows) and geopotential height (contours – 0.1 spacing, solid for 0 and positive, dashed for negative), with topography shown. The dot in each plot shows the location of the Karakoram. Temperature (wind) correlations are shaded (black) only if significant at the 95% level.

To quantify inter-annual variation in the position and intensity of the SWJ, the Karakoram Zonal Index (KZI) is defined. This is calculated by standardising a KZS time series accounting for its seasonal cycle, typically on a monthly or seasonal basis. For example, for a given season, a KZI time series is calculated by extracting all values in the time series for the season, subtracting the season mean and dividing by the season standard deviation. The KZI therefore provides a dimensionless quantification of SWJ position-intensity anomalies relative to its climatology. In all seasons, positive (negative) KZI phases denote northward (southward) position-intensity anomalies relative to the mean state.

Forsythe et al. (2017) and Li et al. (2018) demonstrate that SWJ position-intensity anomalies described by the KZI are associated with characteristic anomalous circulation and temperature responses. These responses form the essence of the KV. For winter and summer, Figure 6.1b shows the KV structure as maps of KZI correlations with 500 hPa horizontal wind component anomalies and temperature. In both seasons, the wind correlation vectors show an anomalous anticyclonic pattern surrounding positive temperature correlations. Therefore, for both seasons, when KZI is positive and the SWJ adopts an anomalous northward position-intensity, the KV shows an anticyclonic wind anomaly that is associated with warmer temperatures. Conversely, negative KZI is associated with a cyclonic wind anomaly and cooler temperatures. Crucially, the area occupied by the KV, especially its longitudinal extent, contracts from a wide expanse in winter to a much more restricted domain centred on the Karakoram in summer.

Moreover, the deep KV system is coupled to the SASM circulation in summer by its southern branch. As such, the KV reflects the interaction between mid-latitude westerly circulation linked to the SWJ and the SASM. KZI can thus be interpreted as a metric of their interplay, as demonstrated by its correlation with monsoonal indices, such as the Webster Yang Monsoon Index (WYMI) (Forsythe et al. 2017).

The connection between the KV circulation and temperature anomalies is considered to be closely related to adiabatic processes, especially in the vicinity of the Karakoram, near the centre of the KV. This is reflected in the latitude-elevation transects for winter and summer in Figure 6.1c. The correlation vectors in this case are for meridional and vertical wind components, with the downward vectors over the Karakoram showing that positive (negative) KZI is associated with anomalous sinking (rising) and thus warming (cooling). These transects also highlight the deep structure of the KV. Importantly, areas of positive

correlations between KZI and temperature extend down to the surface, but in summer the area is much more restricted and centred on the Karakoram. Other influences on temperature anomalies are correlated with KZI states, such as increased cloud cover and so decreased insolation under anomalous cyclonic conditions, for example.

As noted in Section 6.1, Forsythe et al. (2017) argued that the KV likely contributes to the observed summer cooling (Fowler and Archer 2006; Forsythe et al. 2012b), neutral glacier mass balance (e.g. Hewitt 2005; Zhou et al. 2017; Bolch et al. 2017; Brun et al. 2017), and stable or declining river flow trends (Sharif et al. 2013) observed in heavily glaciated sub-basins of the Karakoram and UIB. The mechanism for this is underpinned by the significant negative KZI trend in summer over recent decades, which suggests a southward shift of the SWJ coupled to observed weakening of the SASM (e.g. Saha et al. 2014) and so a higher prevalence of anomalous cyclonic conditions over the Karakoram. This leads in turn to cooling through adiabatic rising and potentially other influences, such as increased cloud cover. This provides climatic conditions conducive to reduced ablation, hence potentially helping to explain observed glacier mass balance and river flow trends, although the connections between air temperature and glacier melt ideally need to be interpreted with reference to the full surface energy and mass balances. Nevertheless, the summer contraction of the KV, such that positive correlations between KZI and temperature are restricted to the Karakoram, appears to provide a strong contribution in explaining why other parts of the Himalayan arc are experiencing rising temperatures and often substantial glacier retreat.

6.3 Data and Methods

6.3.1 CMIP5

This study uses GCM outputs from CMIP5 (Taylor et al. 2012). Spanning a range of modelling groups, configurations and resolutions, the 22 models considered are listed in Table 6.1. Details of the different models are summarised in Chapter 9 (Flato et al. 2013) of the Fifth Assessment Report (AR5) of the Intergovernmental Panel on Climate Change (IPCC; IPCC 2013). To evaluate representation of the KV, we use model outputs from both historical model runs and atmosphere-only (AMIP) simulations. The latter are constrained by realistic sea surface temperatures (SSTs) and sea ice extents. This enables us to detect whether a model's representation of ocean-atmosphere feedbacks play an appreciable role in shaping how its atmospheric component simulates the KV. In the analysis, the historical coupled atmosphere-ocean model runs are referred to with the abbreviation "CMIP", while the

atmosphere-only runs are referred to as "AMIP". These model results underpin the first and second objectives of this work, namely assessing how well the basic KV structure is represented in current GCMs, as well as the connections between the KV and temperature variability and trends in the Karakoram.

In addition, the study uses model outputs from the midrange and high RCP emissions scenarios, RCP4.5 and RCP8.5, respectively. These datasets are used to address the third objective regarding projected changes. Specifically, we use these scenario runs to assess whether the associations between the KV and patterns of near-surface air temperature change may differ in the future. We also consider how large a role the KV has in shaping ensemble spread in projected temperature changes in the Karakoram, as well as the contribution of selected other factors.

The analyses in this study are underpinned by KZS time series (Section 6.2). To calculate these, we extracted time series of zonal winds at 200 hPa spatially averaged over the north and south component areas of the KZS (Figure 6.1), as per Forsythe et al. (2017). KZS and KZI were then calculated on monthly and seasonal timescales. The KZS climatology for historical (CMIP) and AMIP runs can be compared directly to the reanalysis reference datasets introduced in Section 6.3.2. The KZI time series is primarily used as a basis for correlation analysis, with the resulting correlations amenable to comparison with reference data. To elucidate the basic structure of the KV in both horizontal and vertical planes, we calculated maps and latitude-elevation transects of KZI correlations with air temperature, geopotential height and wind components at multiple levels. This correlation-based approach reveals the KV structure in reanalyses, with composite analysis producing similar results (Forsythe et al. 2017; Li et al. 2018). Bilinear interpolation was used to regrid each model to a common grid for comparison with each other and reference data.

In addition, we extracted time series of near-surface and 500 hPa air temperatures spatially averaged over the north-west UIB (NWUIB) domain, which is considered representative of temperature anomaly time series in the Karakoram (Forsythe et al. 2017). This allows us to evaluate the modulating effect of the KV on air temperatures in the models where its influence should be roughly at a maximum. Temperatures were also extracted for a larger 10° square area centred on the NWUIB, in order to test the sensitivity of correlations to the averaging domain chosen. Other variables related to the surface mass and energy balances

were also extracted for the NWUIB domain, particularly to explore how influences on near-surface air temperatures other than the KV are represented in GCMs.

Model	Institution	Resolution (lon° x lat°)
CCSM4	NCAR	1.25 x 0.9
CNRM-CM5	CNRM-CERFACS	1.40625 x 1.40625
MIROC5	MIROC	1.40625 x 1.40625
ACCESS1-0	CSIRO-BOM	1.875 x 1.25
HadGEM2-AO	NIMR-KMA	1.875 x 1.25
HadGEM2-CC	MOHC	1.875 x 1.25
HadGEM2-ES	MOHC	1.875 x 1.25
IPSL-CM5A-MR	IPSL	2.5 x 1.25
inmcm4	INM	2 x 1.5
MPI-ESM-LR	MPIM	1.8 x 1.8
CESM1-CAM5-1-FV2	NSF-DOE-NCAR	2 x 1.9
NorESM1-M	NCC	2.5 x 1.9
GFDL-CM3	GFDL	2.5 x 2
GFDL-ESM2G	GFDL	2.5 x 2
GISS-E2-R	NASA/GISS	2.5 x 2
IPSL-CM5A-LR	IPSL	3.75 x 1.89
FGOALS-s2	IAP	1.66 x 2.81
MIROC-ESM	MIROC	2.8125 x 2.8125
BCC-CSM1-1	BCC	2.8125 x 2.8125
BNU-ESM	GCESS-BNU	2.8125 x 2.8125
CanESM2	CCCMA	2.8125 x 2.8125
FGOALS-g2	LASG-CESS	2.8125 x 2.8125

Table 6.1 – CMIP5 models used in this study.

To assess projected changes, we calculated time series of global near-surface air temperatures using latitude cosine weighting. From this, we identified global temperature changes associated with specific time periods, using 1850-1900 as a pre-industrial reference. We also identified the time periods associated with specific global temperature changes relative to the pre-industrial mean. Using a 30-year moving window, we found the central year where the global temperature change was most similar to the specified target. We use this approach to account at least partially for the variations of equilibrium and transient climate sensitivities between models. This approach allows us to explore the residual spread in regional temperature projections in the model ensemble, but it is subject to limitations of climate responses not being stabilised (James et al. 2017). This is discussed further in conjunction with the results.

6.3.2 Reference Datasets

The main reference datasets used for evaluation of the CMIP5 models here are global reanalysis products. Specifically, the ensemble used comprises ERA-40 (Uppala et al. 2005), ERA-Interim (ERAI) (Dee et al. 2011), JRA-55 (Kobayashi et al. 2015), MERRA (Rienecker et al. 2011) and CFSR (Saha et al. 2010). As noted above, the KV and its relationship with temperature have been shown to exist in these reanalyses (Forsythe et al. 2017; Li et al. 2018). The reanalyses show strong similarity in their representations of the KZS cycle, as well as the sequencing of inter-annual KZI anomalies over recent decades. All of the reanalyses also show similar patterns of correlation between KZI and temperature, albeit with some variation in the magnitude of correlation. As such, we use the full reanalysis ensemble to evaluate the representation of KZS and KZI correlations with temperature. For looking at maps and transects of the basic KV structure, we focus on ERAI as a representative data product, based on Li et al. (2018).

As additional reference data, we also use long-term locally observed temperature records from the Astore, Gilgit and Skardu stations run by the Pakistan Meteorological Department (PMD). These data have been subjected to quality control and extensive analysis previously (e.g. Fowler and Archer 2006; Forsythe et al. 2017). Data for the period 1958-2005 are largely complete and are primarily used here as an additional reference in trend analysis. Previous work has demonstrated the substantial skill of the reanalysis ensemble in capturing the sequencing of inter-annual temperature anomalies, such that we rely on reanalyses for producing the reference correlation analyses rather than local observations (Forsythe et al.

2017). The Collection 5 MODIS MCD43A3 data product is used as an additional reference dataset for albedo.

6.4 Results

6.4.1 Karakoram Vortex (KV) Representation

Karakoram Zonal Shear (KZS)

To evaluate how well CMIP5 GCMs simulate the KV, we first examine their representation of the KZS. Figure 6.2a shows that the annual cycle of the KZS is essentially well reproduced by the GCM ensemble. Some degree of spread is evident, but in general the shape, timing and amplitude of the cycle are largely consistent with reanalysis. Crucially, this indicates that the seasonal migration of the SWJ is reasonably simulated at the ensemble level. However, there is some tendency to marginally under-predict KZS, most notably in winter but also in spring and summer to varying degrees, although not all models exhibit this issue. Figure 6.2b shows that overall this under-prediction bias is slightly lesser in the atmosphere-only (AMIP) model ensemble compared with the historical (CMIP) runs. This suggests that bias in the coupled model ensemble could be related to deficiencies in sea surface temperature (SST) or sea ice extent in coupled model simulations. Nevertheless, overall KZS biases are essentially small relative to the amplitude of the annual cycle.

Regional Temperature and Wind Correlations

To assess whether reasonable KZS performance translates into the correct horizontal structure of the KV, maps of correlations between KZI and 500 hPa temperature (T500) and wind anomalies are introduced. The maps of winter (DJF) correlations in Figure 6.3 demonstrate that the basic traits of the KV are present in the CMIP ensemble at this time of year. Focusing on temperature first, a broad longitudinal band of positive KZI/T500 correlations is clearly in evidence in the ensemble, in agreement with the ERAI reference. The highest correlation coefficients are also correctly located in the vicinity of the Karakoram. Areas of significant KZI/T500 correlation in the multi-model mean (MMM) do not extend as far eastwards as in ERAI, although the MMM does reproduce both the latitudinal range and the westerly extent of significant correlations. Correlation strength is slightly lower in the MMM than ERAI, but amongst individual models there is substantial variation in the magnitude, areal extent and shape of zones with significant correlations. Some models have larger areas of strong positive

correlation than in ERAI (e.g. CanESM2, CCSM4), while others show much more modest positive correlation extents (e.g. CNRM-CM5 and GFDL-CM3).

Figure 6.3 also confirms that in winter the CMIP models generally reproduce the anticyclonic pattern of KZI/wind correlation vectors shown for ERAI. The relationship between positive (negative) KZI, anticyclonic (cyclonic) circulation anomalies and positive (negative) temperature anomalies that characterises the KV is therefore present here. Again, inter-model variation in correlations is notable. Comparing individual models shows a generally close connection between the areal extent and shape of KZI/T500 and KZI/wind correlations. Models with larger bands of positive KZI/T500 correlation tend to show larger anticyclonic structures for KZI/wind correlation vectors and vice versa. While the anomalous circulation response to KZI is therefore present overall, some notable inter-model differences exist in its size, centroid and shape that are associated with the temperature response.

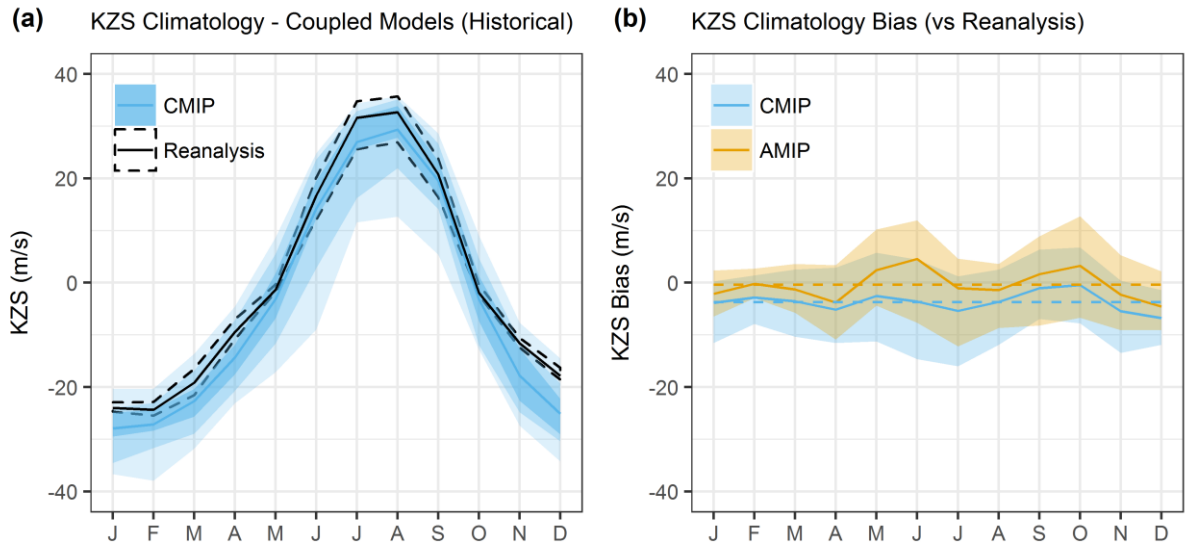


Figure 6.2 – Representation of the Karakoram Zonal Shear (KZS) annual cycle in CMIP5 GCMs. Coupled model monthly climatology (blue) compared with the reanalysis ensemble (black) is shown in (a). Monthly range, 10^{th}-90^{th} and 25^{th}-75^{th} percentiles are shown with successive shading for the GCMs, and only the range is shown for the reanalysis ensemble. Lines represent ensemble means. In (b), biases in the KZS climatology for the coupled historical (CMIP) and atmosphere-only (AMIP) model ensembles are shown. Lines denote ensemble mean biases, while shading shows only the 10^{th}-90^{th} percentiles for clarity.

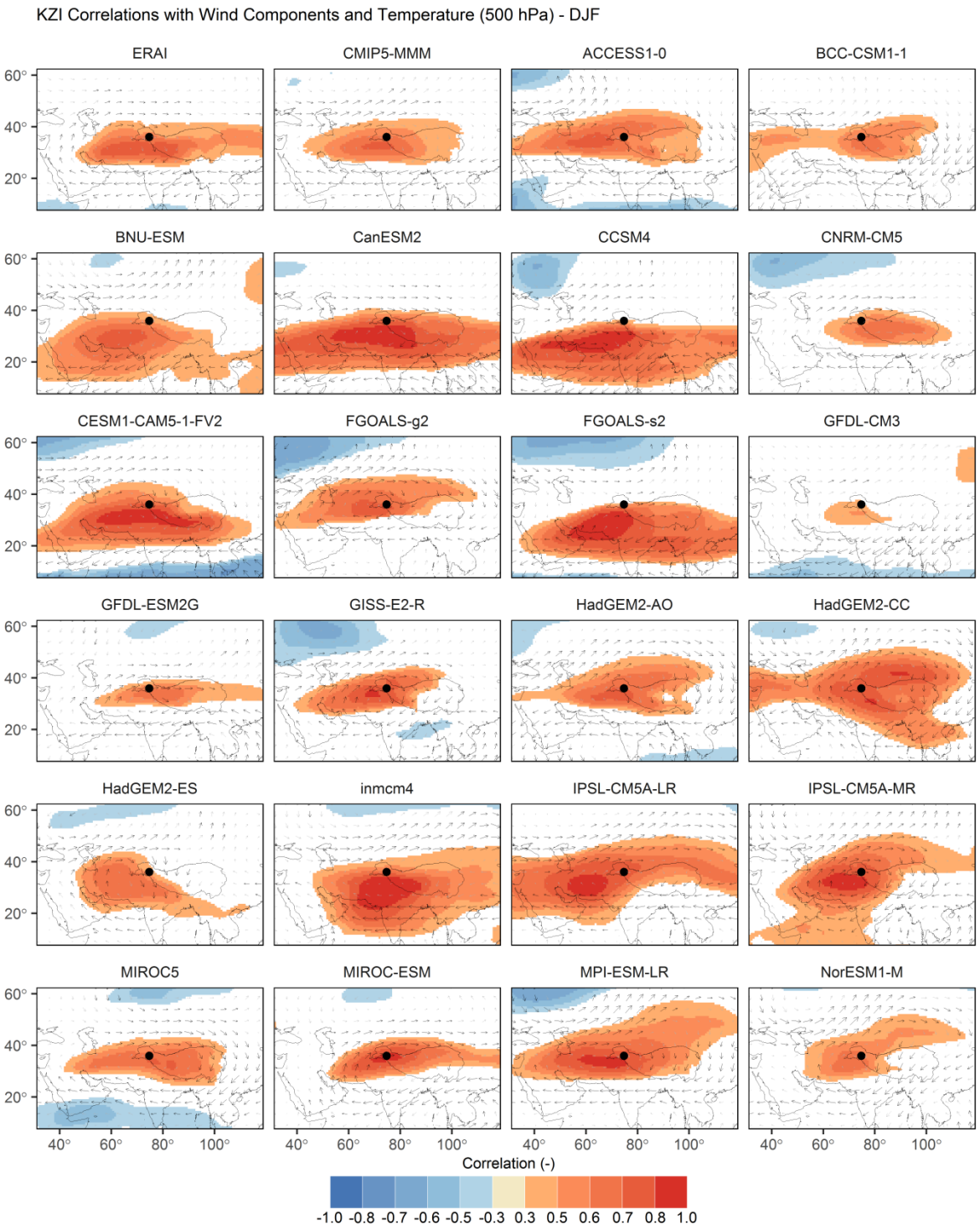


Figure 6.3 – Winter (DJF) correlation of KZI with temperature (shading) and horizontal wind components (arrows) at 500 hPa in CMIP5 compared with ERAI (1979-2005). Temperature correlations are shown if significant at the 95% level. Wind correlations are shown in black if either component is significant at the 95% level and in light grey otherwise. The dot marks the Karakoram.

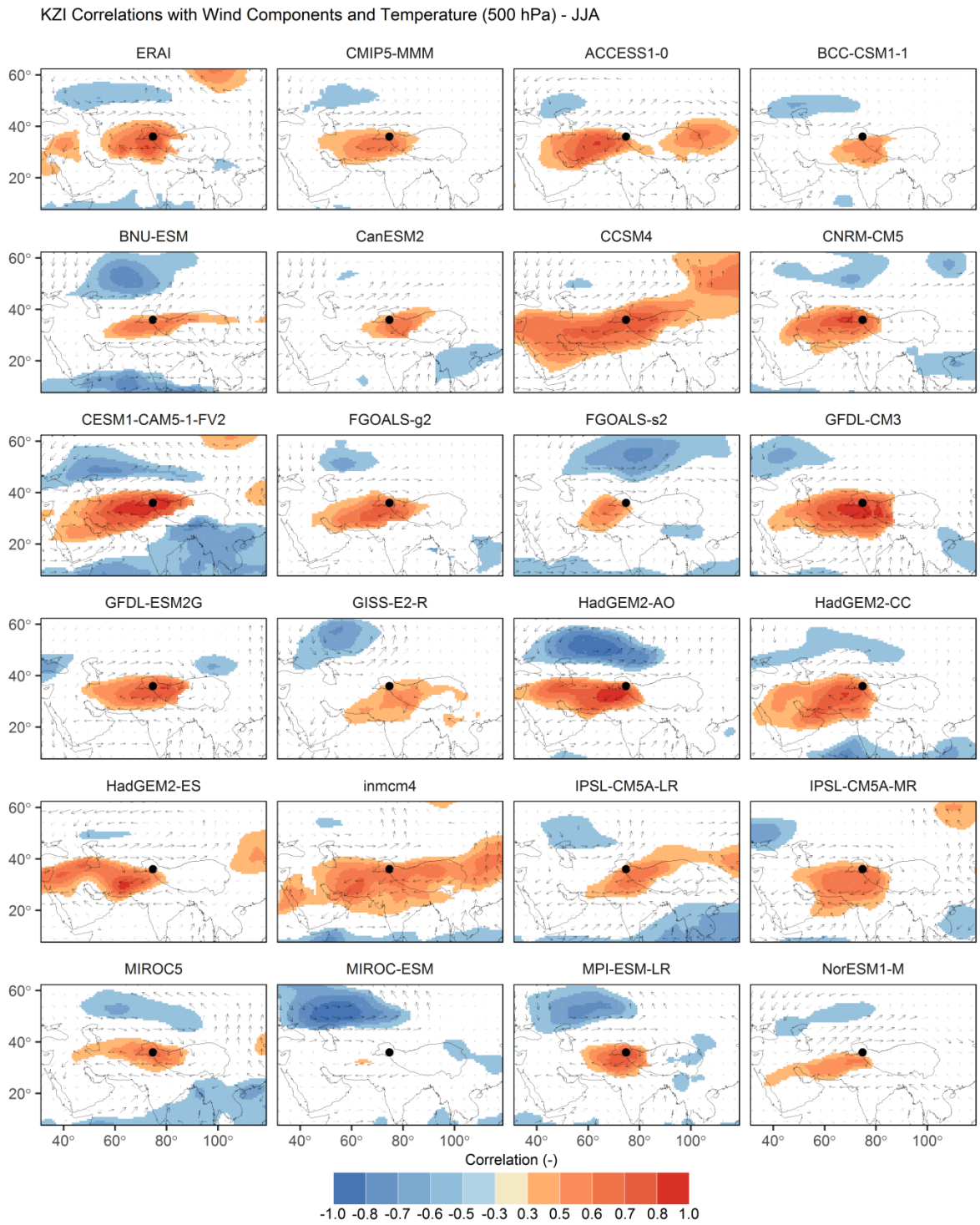


Figure 6.4 – As Figure 6.3 but for summer (JJA).

For summer (JJA), Figure 6.4 demonstrates that the most of the CMIP models capture the contraction of the band of positive KZI/T500 correlations relative to their wider extent in winter. The positive correlations become largely restricted to the Karakoram and the western Tibetan Plateau, with the anticyclonic structure of the KZI/wind vectors generally contracting

accordingly. This is well represented in the MMM when compared with ERAI. The degree of contraction of the KV influence area differs notably between individual models, with some models retaining overly large zones of positive correlation (e.g. CCSM4, inmcm4). In some cases, the contraction of the KZI/wind correlation vectors is inadequate, with the location and/or coherence of the KV structure becoming unclear. A centre of negative KZI/T500 correlation corresponding with a cyclonic pattern of KZI/wind correlation vectors is also present in most models to the north of the zone of positive correlation characterising the KV. However, the correlation strength and size of this centre vary substantially between models.

The CMIP5 models can thus essentially reproduce the key features of the KV in terms of temperature and wind at 500 hPa. We now ask whether they capture the relationship with near-surface (2m) air temperature T2. This is critical, as the KV is known to provide a substantial modulating influence close to the surface, which is often the key focus for many glaciological and hydrological applications of GCMs. For winter, Figure 6.5 indicates that, while a broad band of significant positive KZI/T2 correlation exists in ERAI, the coherence of this band breaks down in the CMIP ensemble to varying degrees. The MMM shows positive correlation over a fairly restricted area compared with ERAI and the 500 hPa case (Figure 6.3). Correlations in the MMM are not statistically significant over much of the interior of the Tibetan Plateau. Indeed, a number of the individual models show breaks in the significance of their correlations that follow the shape of the TP or parts of it. This suggests some influence of the Tibetan Plateau's topography and near-surface processes on temperature variability, diminishing the more extensive modulating influence of the KV at the 500 hPa level.

With contraction of the KV in summer, the significant positive KZI/T2 correlation becomes largely restricted to the Karakoram and western Tibetan Plateau. This is accompanied by a zone of negative KZI/T2 correlation to the south (over the Indo-Gangetic plain) and extending eastward. Figure 6.6 shows that, compared with ERAI, the MMM exhibits these basic patterns but with notably smaller areal extents of significant correlations. Again, there is appreciable spread when examining individual models. Most models do capture the pattern of positive correlations in the Karakoram and western Tibetan Plateau but negative correlations to the south and east. Many of the models appear to simulate relatively small zones of positive correlation, with the exception of the overly large extent in CCSM4. However, some models do not appear to exhibit very much significant correlation over the Karakoram and western Tibetan Plateau at all.

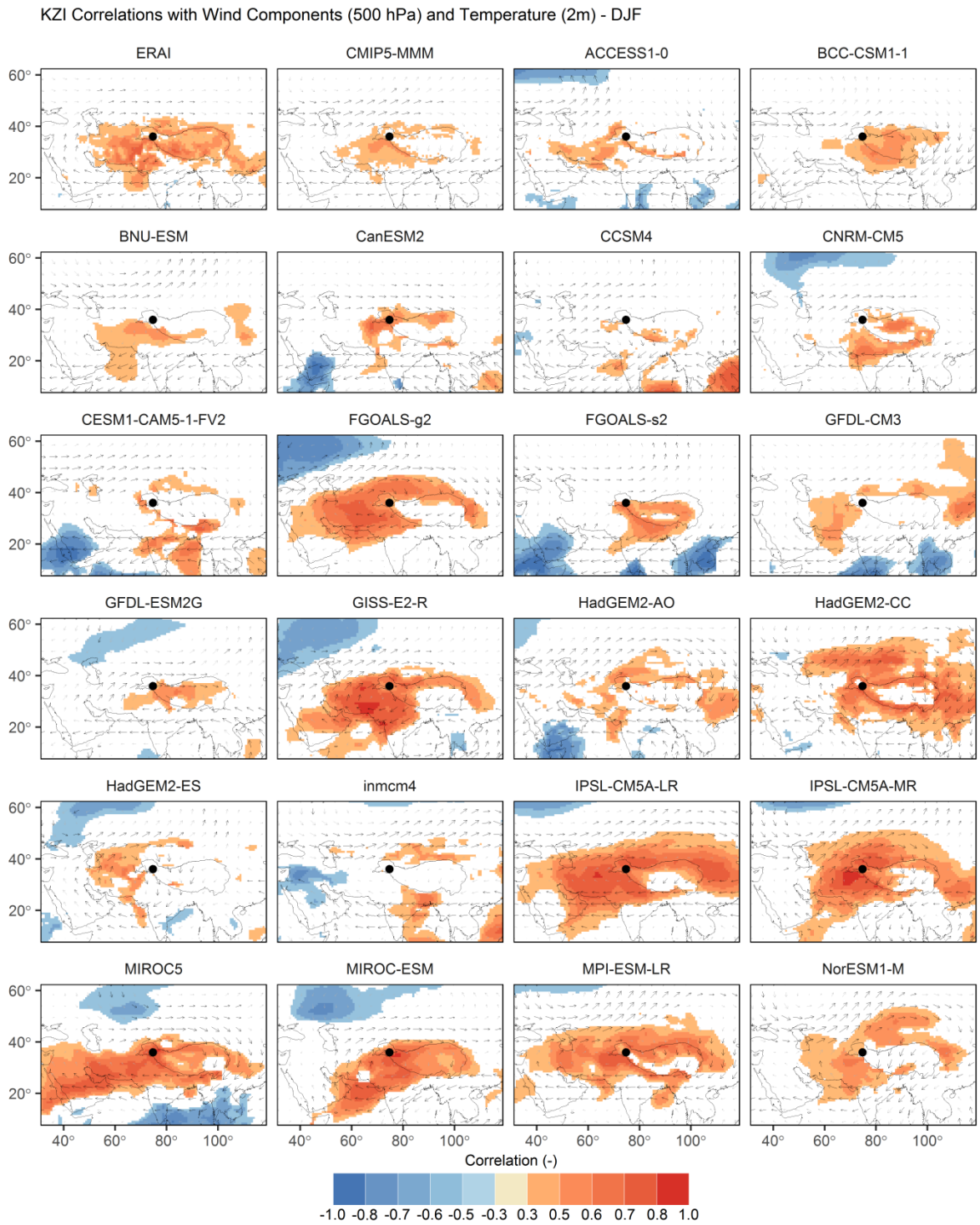


Figure 6.5 – Winter (DJF) correlations as Figure 6.3 but showing 2m air temperature (rather than 500 hPa temperature).

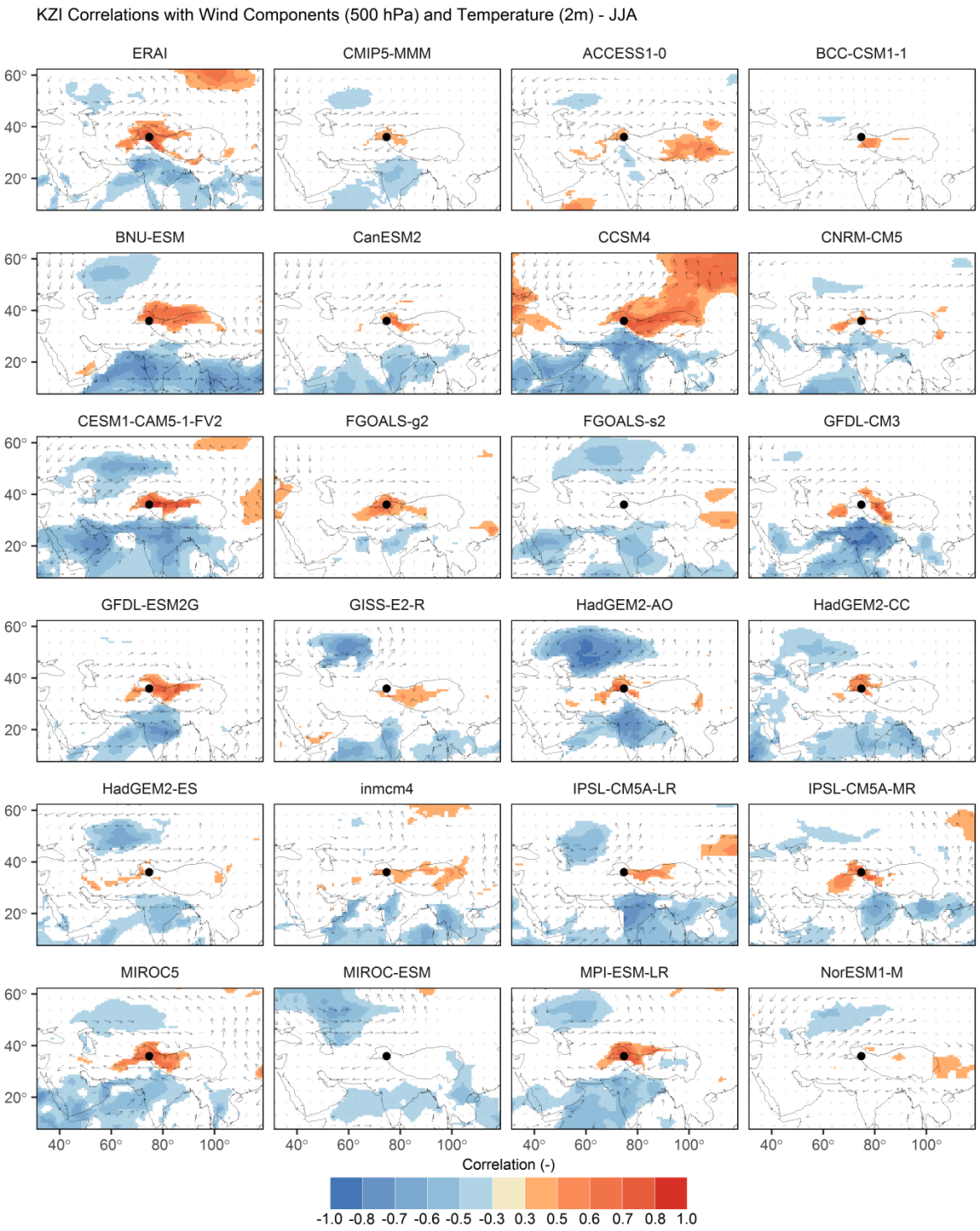


Figure 6.6 – Summer (JJA) correlations as Figure 6.4 but showing 2m air temperature (rather than 500 hPa temperature).

Latitude-Elevation Transects

The basic KV structure can be further assessed with respect to latitude-elevation transects. Figure 6.7 shows these transects passing through the Karakoram for winter. From this, we can see that the CMIP models agree with ERAI on a number of key aspects of the vertical structure of the KV. Crucially, the vertical dipole of positive and negative correlations between KZI and temperature is present in virtually all models. Both latitudinal and vertical variations in correlation are well represented in the MMM, although the strength of correlation is again a little lower than in ERAI. As inferred from the maps in Figure 6.3, some models show much larger latitudinal extents of significant positive correlation; Figure 6.7 indicates that these relatively large correlation extents may be present on multiple pressure levels, up to the change in correlation sign in some cases. In addition, many of the CMIP models show the zone of positive correlation extending downwards to the surface. While Figure 6.5 illustrated relatively large patchiness in (significant) KZI/T2 correlations, averaging over a longitudinal band confirms that there is a tendency for positive correlations to exist down to the surface in most models.

For models where KZI is significantly correlated with the meridional (v) and vertical (w) wind components over the Karakoram, the correlation vectors are in broad agreement with ERAI. Specifically, the correlation vectors point downwards, and to a lesser degree southwards. The downwards component indicates an increase in vertical wind velocity towards the surface when KZI increases and vice versa. This is consistent with ERAI and the inference of adiabatic warming (cooling) associated with the anomalous anticyclonic (cyclonic) circulation patterns that characterises positive (negative) KZI states. However, not all models exhibit significant KZI/wind correlation vectors over the Karakoram. This diversity means that the MMM also does not contain significant correlations in this respect.

In summer, the CMIP models generally simulate the latitudinal contraction of the KV, in terms of the positive correlations between KZI and temperature (Figure 6.8). The models also capture the retreat of the zone of positive correlations upwards away from the surface, giving way to negative correlations to the south and north in most models. Acknowledging inter-model variation, this critical feature of the KV annual cycle is therefore present. Similar to the winter case, the KZI/wind correlation vectors point downwards over the Karakoram in ERAI and a number of the models. However, not all models display significant correlations in this respect.

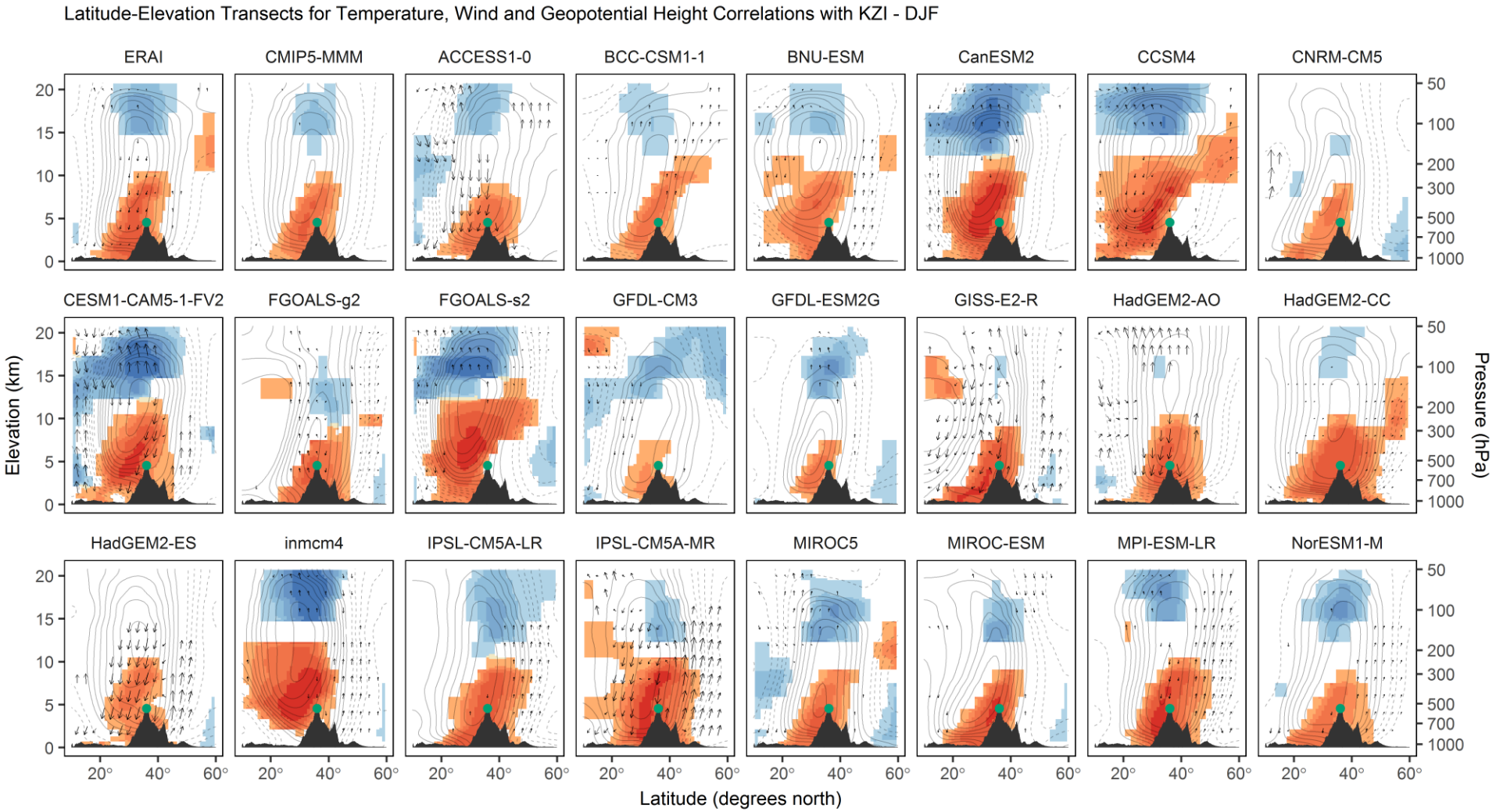


Figure 6.7 – Latitude-elevation transects (integrated over 70-80°E) of winter KZI correlations with temperature (shading), meridional and vertical wind (arrows) and geopotential height (contours – 0.1 spacing, solid for 0 and positive, dashed for negative), with topography shown. Correlation scales are as Figure 6.3.

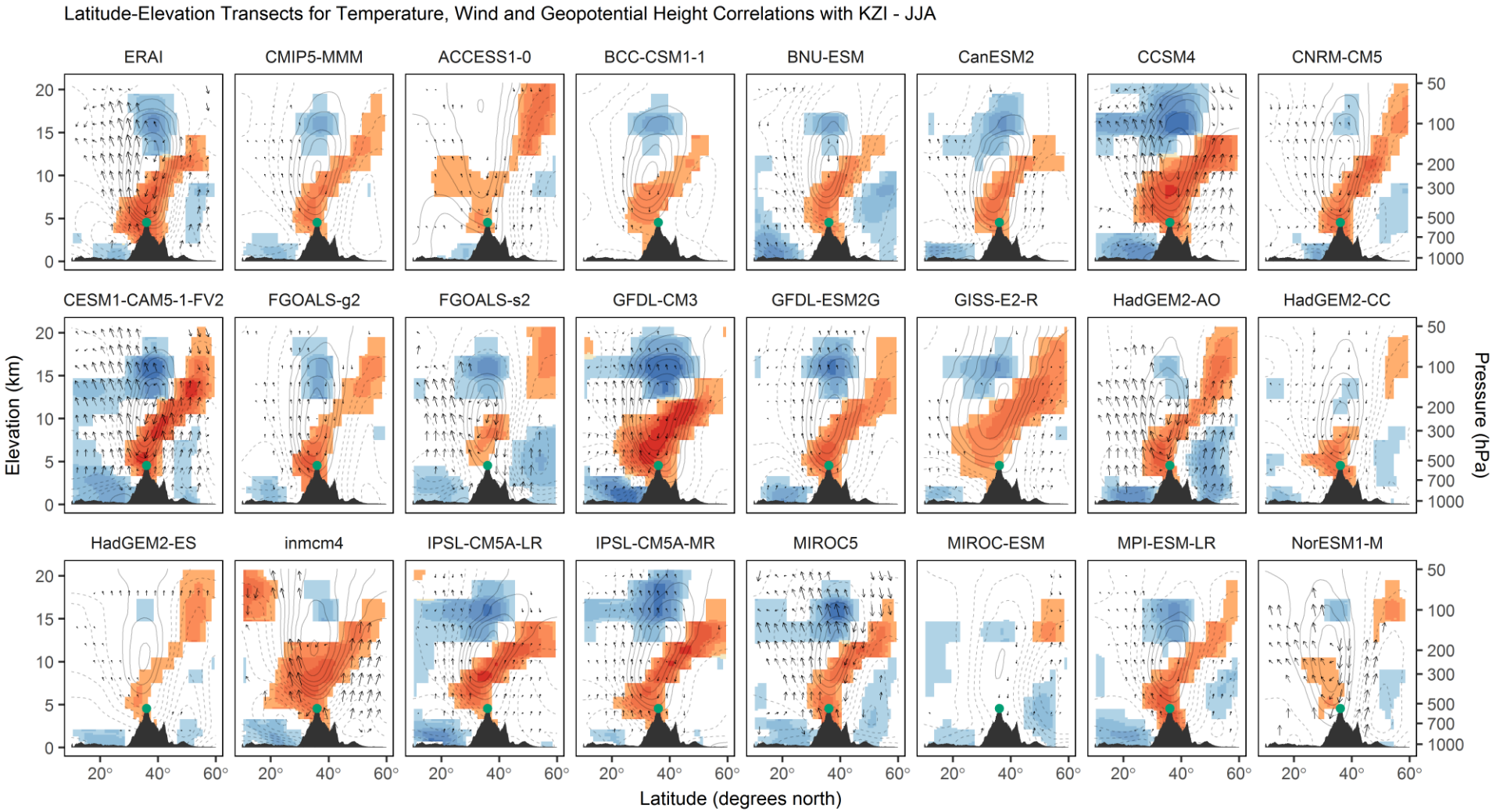


Figure 6.8 – As Figure 6.7 but for summer.

6.4.2 Karakoram Inter-Annual Temperature Response

Section 6.4.1 thus demonstrates that the key circulation and temperature responses associated with the KV are often present in GCMs, but there is notable inter-model variation and potential deficiencies in the strength of correlations with inter-annual KZI anomalies. We now focus in on the Karakoram region to explore this in more detail.

Figure 6.9a compares KZI correlations with NWUIB T2 in the CMIP ensemble with the reanalysis ensemble described in Section 6.3.2. This confirms that the positive relationship between KZI and T2 throughout the year is present in most CMIP models. However, the ensemble mean (and inter-quartile range) tends to be generally lower than the reanalysis mean for much of the year. This is in agreement with the spatial patterns of correlation in Figure 6.5 and Figure 6.6. Interestingly, correlations in late autumn and early winter (October to December) do not experience the same reduction as is evident in the reanalysis ensemble. Rather, CMIP correlations are fairly consistent in all seasons. Notably, the bands showing the 10^{th}-90^{th} percentiles and the overall range on Figure 6.9a demonstrate that at least some models exhibit very weak KZI/T2 correlations over the Karakoram, while the models showing the strongest correlations reach the same levels as the reanalysis maxima (indeed exceeding them between October and December).

Figure 6.9a therefore indicates that the distribution of KZI/T2 correlation coefficients within the CMIP ensemble is essentially characterised by a general low bias relative to reanalyses in most of the year for the NWUIB, but also notable inter-model variability. The differences between models are shown in full in Appendix C (Figure C1). One point to note from this is that 95% confidence intervals for the KZI/T2 correlations overlap with the range of reanalysis correlations for many models/months. While it is still likely that a low bias exists at the ensemble level, this highlights that differences for individual models/months should be interpreted with care. To evaluate the sensitivity of the correlations to the box used for T2, it is also demonstrated in Appendix C that the results are very similar if a 10° box centred on the Karakoram is used in the correlations, rather than the NWUIB domain (Figure C2). There are strong linear relationships in all seasons between KZI/T2 correlations using these two domains for T2, although the degree of scatter is noteworthy.

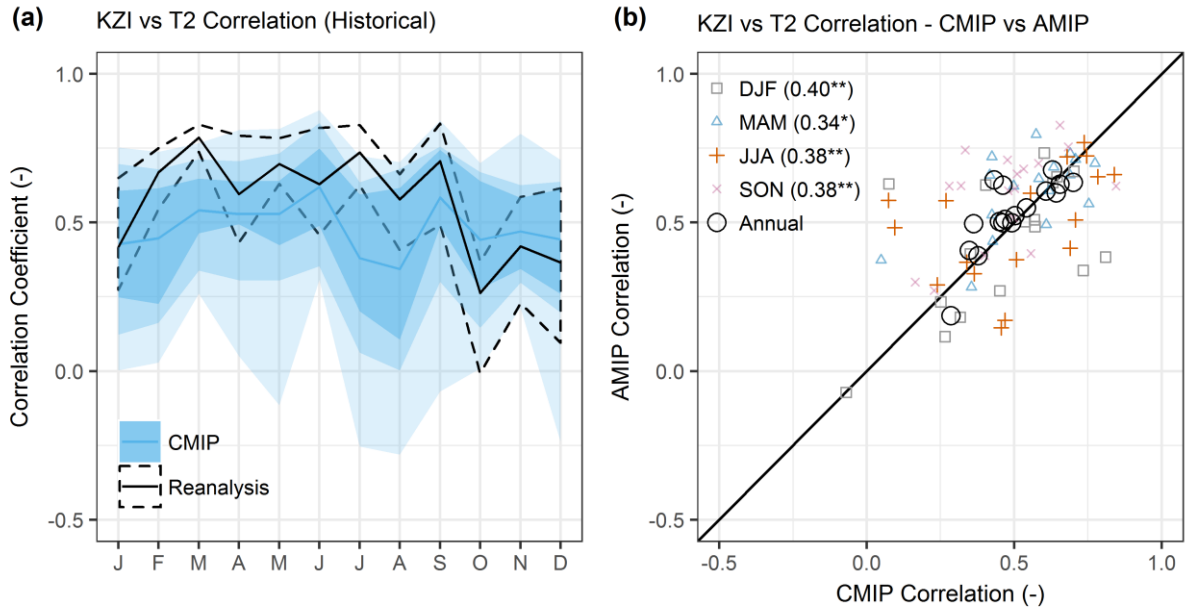


Figure 6.9 – (a) Monthly correlation of KZI and T2 in the NWUIB in CMIP5 and ERAI (1979-2005). The mean and range are shown for the reanalysis ensemble in black. The mean, 25-75th, 10-90th and total ranges are shown in successive blue shading for the CMIP5 ensemble. (b) Scatter plot of seasonal correlations between CMIP and AMIP scenarios for individual models. Correlation coefficients (Kendall's tau) are given in brackets, with significance at the 90 and 95% levels shown by * and **, respectively.

A key question arising from this is what might cause the inter-model variation. Beginning with broader scale influences, we consider whether differences between CMIP and AMIP model runs help to explain inter-model differences in KZI/T2 correlation strength. Figure 6.9b shows the relationship between correlations in the CMIP and AMIP ensembles, with the latter prescribing observed SSTs and sea ice to reduce key sources of bias in coupled atmosphere-ocean model simulations (Section 6.3.1). Essentially this shows that KZI/T2 correlations in CMIP and AMIP are positively related, falling around the 1:1 line. This means that, on average, KZI/T2 correlations in any given model are approximately equal in the CMIP and AMIP experiments. There is deviation present, but overall this scatter appears to be largely unstructured, possibly due to natural variability. The conclusion here is then that inter-model variations in simulated KZI/T2 correlations in the CMIP ensemble overall are not particularly related to any biases stemming from SST or sea ice errors. As shown in Figure 6.1b, some such biases influencing regional circulation (quantified by KZS) may be present in the CMIP ensemble that are reduced in the AMIP ensemble, but these biases appear to be fairly small and not closely linked to strength or biases in KZI/T2 correlation for the ensemble as a whole.

Model	KZI vs T2 Correlation		KZS		T2	
	RMSE (-)	Rank	RMSE (m/s)	Rank	RMSE (°C)	Rank
CESM1-CAM5-1-FV2	0.16	1	4.90	7	4.49	15
GFDL-ESM2G	0.17	2	5.23	9	1.12	1
HadGEM2-CC	0.17	3	6.94	13	2.49	4
MPI-ESM-LR	0.20	4	4.45	4	2.52	5
FGOALS-g2	0.21	5	7.55	16	2.25	2
ACCESS1-0	0.21	6	6.62	11	3.93	11
IPSL-CM5A-MR	0.21	7	7.92	18	4.28	13
MIROC5	0.21	8	3.99	3	3.80	10
HadGEM2-AO	0.22	9	6.63	12	3.12	7
BCC-CSM1-1	0.23	10	7.61	17	3.78	9
CCSM4	0.23	11	7.15	15	5.41	20
MIROC-ESM	0.25	12	3.77	2	3.17	8
GISS-E2-R	0.25	13	10.57	21	4.37	14
CanESM2	0.26	14	4.48	5	2.91	6
HadGEM2-ES	0.26	15	6.54	10	2.29	3
IPSL-CM5A-LR	0.27	16	11.54	22	5.21	18
BNU-ESM	0.29	17	7.01	14	4.94	16
NorESM1-M	0.35	18	9.53	19	4.27	12
FGOALS-s2	0.36	19	4.92	8	5.26	19
CNRM-CM5	0.38	20	4.77	6	11.43	22
GFDL-CM3	0.38	21	10.50	20	8.42	21
inmcm4	0.53	22	3.19	1	5.06	17

Table 6.2 – CMIP model performance.

Another possibility is that inter-model in KZI/T2 correlation is related to KZS or T2 biases or errors. Table 6.2 summarises individual model performance for the CMIP runs, which indicates that the relationship between these three criteria are not straight-forward. It is possible to show good performance for one criterion and bad performance for at least one other, and vice versa. The point is further demonstrated in Appendix C, where Figure C3 and Table C1 additionally quantify the relationships between bias in KZS, T2 and KZI/T2 correlations on a seasonal basis. This suggests that the relationships are generally weak. There may be some more connections in winter, but the scatter appears to be very substantial. As such, it is suggested that the relationships between KZS, T2 and KZI/T2 performance are somewhat complicated.

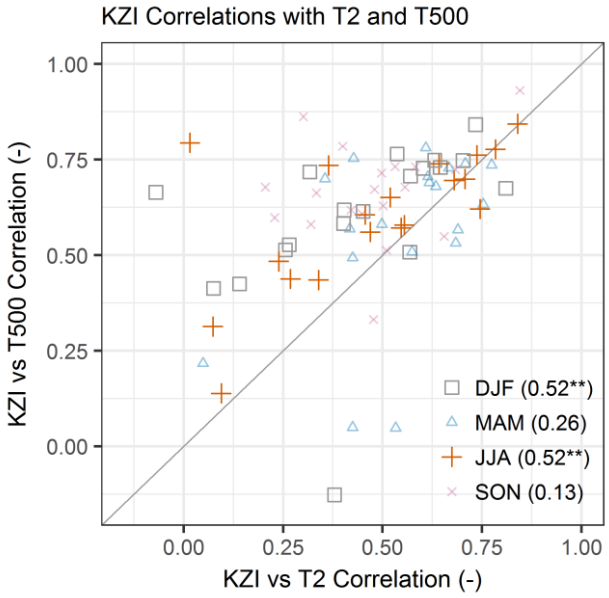


Figure 6.10 – Comparison of seasonal correlations between KZI and temperature at 2m (T2) and 500 hPa (T500) for CMIP models. Correlation coefficients (Kendall's tau) are given in brackets, with significance at the 90 and 95% levels shown by * and **, respectively.

One possibility for this is that the KZS provides an incomplete characterisation of the KV system and related processes. Another is that other aspects of model behaviour compromise the modulating influence of the KV on temperatures. To explore this further we compare KZI correlations with temperature for the NWUIB domain at 500 hPa and 2m levels in Figure 6.10. This shows that there are fairly strong, ostensibly linear relationships between KZI and temperature correlations at these levels for winter (DJF) and summer (JJA). Although a reasonable degree of scatter is present, this means that models with higher correlations between KZI and temperature at one level will also have higher correlations at the other in

general. Interestingly, the strength of relationships appears to be lower for spring (MAM) and autumn (SON), which are in effect transition seasons between the two end-member states of the annual KV and temperature cycles.

In all seasons there are more points above the 1:1 line in Figure 6.10 than below it, indicating that correlations between KZI and temperature are higher at 500 hPa than 2m in general. This suggests some decay in KV modulating influence from the mid-lower troposphere to the surface, which could reflect the increasing influence of near-surface processes on temperature variability, diminishing the signal from the larger scale dynamical state of the KV. This is explored in more detail below.

6.4.3 Karakoram Historical Trends

The preceding sections showing that the basic KV structure is reproduced in GCMs overall, while the KV's strong relationship with inter-annual variation in Karakoram temperatures is generally present if slightly weak. This section now considers the relationship between KZI and T2 trends. While inter-annual variability in the CMIP runs is not expected to align with observed anomalies, it seems reasonable to ask whether what (if any) trends are present over the 1958-2005 period of overlap with observations, as well as whether the connection between negative KZI and T2 trends in summer found by Forsythe et al. (2017) is present.

Figure 6.11 shows historical (1958-2005) KZI and T2 trends in the CMIP ensemble, as well as trends from the two reanalyses beginning in 1958 and observed temperatures. For winter, we can see that most models predict relatively limited KZI changes but some T2 increase (Figure 6.11a). This is similar to reanalysis and observations. The confidence intervals for most of the CMIP models in winter overlap with each other and reanalysis/observations, indicating little discernible difference. Figure 6.11a suggests that there may be some negative relationship between the KZI and T2 trends, in line with the association of negative KZI and T2 anomalies, but the scatter is substantial and the significance difficult to gauge. In spring, the CMIP models largely simulate a warming trend of similar magnitude to the observed T2 trend (Figure 6.11b). The negative trend for KZI in the reanalyses is present in some CMIP ensemble members, but more of the models indicate limited trend.

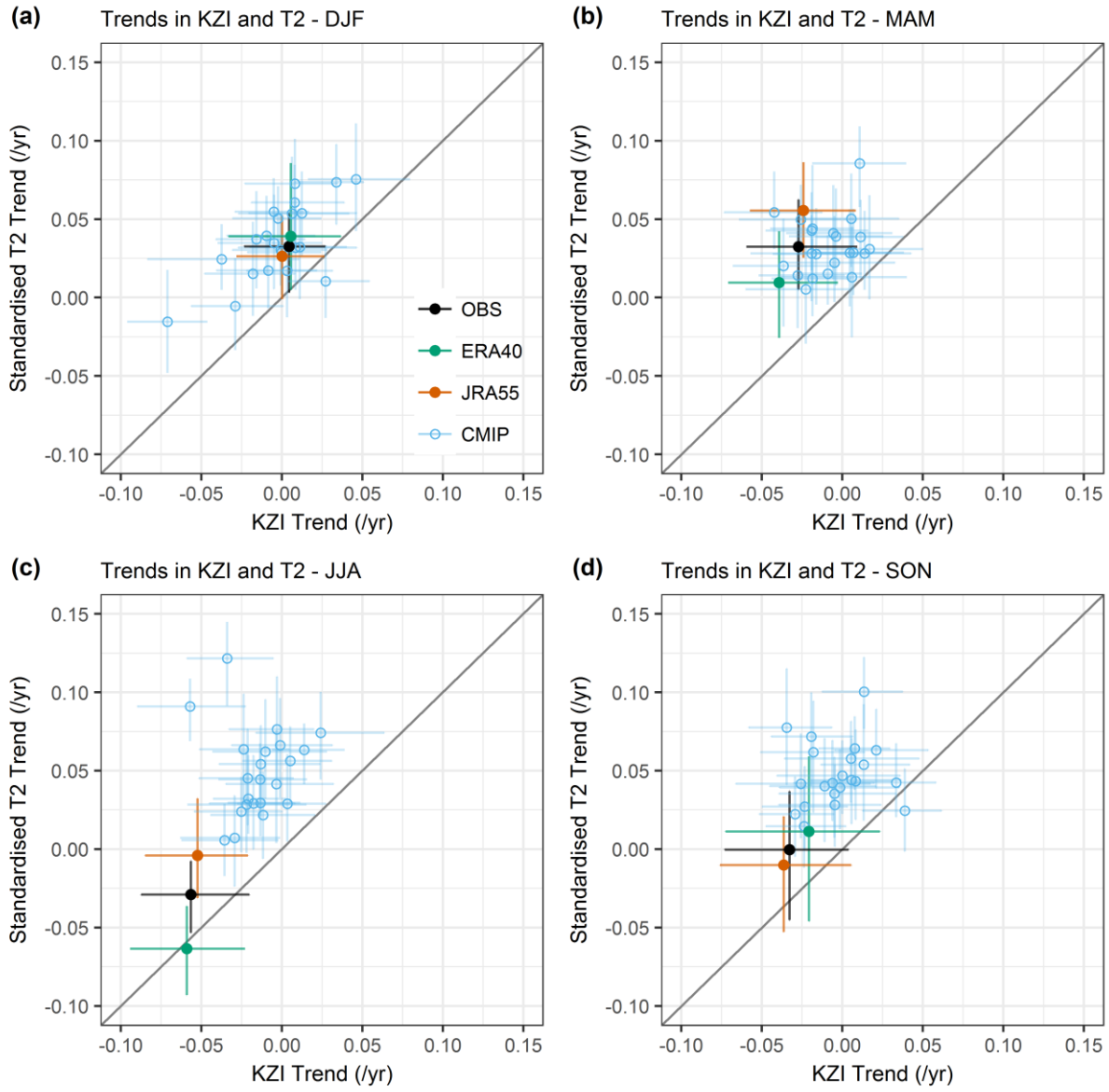


Figure 6.11 – Historical trends in KZI and (standardised) 2m air temperature (T2) by season from 1958-2001. Circles/dots show the trends estimated using the Sen's slope method with the full time series, while lines indicate confidence intervals obtained via block bootstrapping. CMIP models are shown in blue. ERA40 and JRA55 reanalyses are shown; the other reanalyses have shorter record periods. Observed temperature trends are calculated from station observations, with the reanalysis ensemble mean used as a proxy for observed KZI.

More striking differences are found in summer (Figure 6.11c). Here, we can see that the reanalyses show a clear negative trend in historical KZI, which is generally weaker in the CMIP models. Based on the confidence intervals, only a small number of the CMIP models have negative trends that are likely to be significant, but a number of the KZI confidence intervals do overlap with those of the reanalyses. However, while some of the KZI trends in the CMIP ensemble may therefore be in reasonable agreement with the reanalysis margins of

uncertainty, the models diverge substantially from observations in terms of T2 trend. Almost all models predict warming in summer over recent decades rather than the observed cooling. Only two models have confidence intervals overlapping those of the T2 trend in observations; incidentally these models also have negative KZI trends. Indeed, if the two outliers with negative KZI but positive T2 trends are temporarily ignored, visually there is some suggestion of a relatively weak (negative) relationship between KZI and T2 trends in the CMIP ensemble. There are some similarities in autumn with these summer patterns (Figure 6.11d), but in this case the reanalyses show no discernible T2 trend as opposed to cooling. The negative KZI trend is also weaker and the differences from the reanalyses less clear.

A preliminary comparison in Appendix C (Figure C4) suggests that historical summer trends in KZI and T2 in the AMIP ensemble are similar to those in CMIP for their overlapping period of record. In both cases the GCMs show faster warming than observations and the reanalysis ensemble. However, it should be noted that the confidence intervals in both cases (and in Figure 6.11) are fairly large. Moreover, the AMIP runs are comparatively short (1979-2005), such that trend analysis is at best indicative, although the suggestion that CMIP and AMIP temperature trends are less coupled with KZI trends than in observations and reanalyses is intriguing.

The summer discrepancies could indicate that some important circulation connections are poorly simulated, leading to underestimation of the negative KZI trend. There is also the question of whether the summer T2 warming in the CMIP model is too closely connected to other factors, such as the level of "background" global warming or land- and near-surface processes and feedbacks, with the modulating influence of the KV insufficiently captured. Indeed, this would be consistent with the general tendency for the models to underestimate the KZI/T2 correlation (Sections 6.4.1 and 6.4.2). These issues are considered in more detail below.

6.4.4 Projected Changes

The analysis of trends and changes now moves to consider future projections, focusing again on the Karakoram region. We begin by looking at temperature changes expected in the twenty-first century and their relationships with the KV. We then focus on global temperature stabilisation scenarios to explore the residual spread in the CMIP5 ensemble and its drivers.

Twenty-First Century

Figure 6.12 examines the projected changes in T2 and KZS with respect to the accompanying global temperature (T2) change for the future period 2051-2100. The first feature to note is the clear, positive relationship between global and NWUIB T2 changes. In all seasons, T2 warming in the NWUIB tends to exceed the global changes, and by an increasing amount for larger global changes. There is thus amplification of warming in the NWUIB relative to the global signal, but the degree of global change predicted in a given model/scenario generally exerts a strong control on the magnitude of the local response in the NWUIB. However, scatter in the relationship between global and NWUIB T2 changes does vary seasonally; it is at its lowest in winter (Figure 6.12a) and its highest in summer (Figure 6.12c), as confirmed by the correlations in Table 6.3. Therefore, while the global T2 change is a particularly strong predictor of the NWUIB T2 change in winter, the strength of the relationship decreases in summer. As the influence of local scale processes on T2 variability and change is heightened in summer, the larger scatter in Figure 6.12c therefore likely reflects the importance of inter-model variations in land surface process representations and thus land-atmosphere interactions. In contrast, winter T2 changes in the NWUIB appear to be more readily determined from differences in larger scale changes. Spring and autumn appear to essentially represent intermediate cases (Figure 6.12b,d), while all seasons show the highest degree of scatter for the highest global temperature changes.

Moreover, there are generally consistent seasonal differences in the sign of change in KZS accompanying the global and NWUIB T2 changes. In autumn and winter, KZS changes in the future period are more often positive than negative, corresponding with a northward shift in the SWJ and more anomalous anticyclonic conditions over the Karakoram/WTP. However, in spring and summer, most of the model/scenario combinations are associated with negative changes in KZI. This effectively indicates a continuation of the historical pattern, particularly in summer, whereby (potentially underestimated) negative KZI trends are accompanied by T2 warming rather than observed cooling. Interestingly then, the CMIP models project an increase in T2 despite a decrease in KZI. In contrast, reanalyses/observations indicate negative trends in KZI and T2 in summer for the historical period, with the former argued to be a strong contributing driver for the latter.

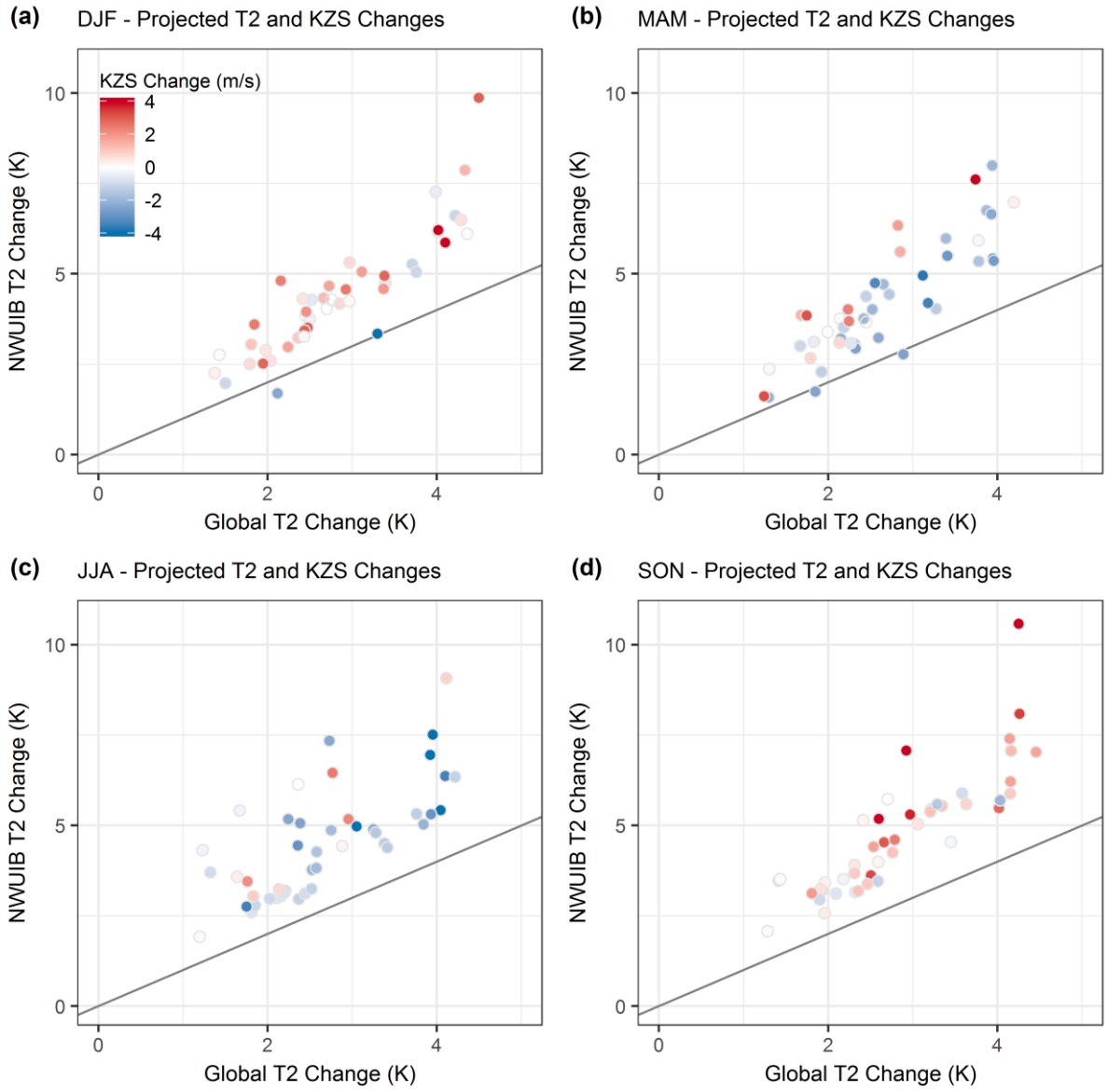


Figure 6.12 – Comparison of projected changes in global T2, NWUIB T2 and KZS for 2051-2100 relative to 1951-2000.

As the CMIP models do not fully capture this feature of the historical climate, it is plausible that they could also misrepresent the future of both the KZI and its modulating influence on T2 variability and potentially change. However, it could also indicate that factors other than the KV become dominant in determining the sign of change, due to "background" global warming and feedbacks or more local processes. We also note that there appears to be generally limited correlation between the magnitude of KZI changes and global or NWUIB T2 changes. This is reflected by the low correlation coefficients in Table 6.3, which suggest some comparatively weak correlation with global T2 change (except winter), as well as NWUIB T2 change in summer and autumn.

Season	Correlation Coefficient (-)		
	Global T2 / NWUIB T2	Global T2 / KZI	KZI / NWUIB T2
DJF	0.76**	0.02	0.11
MAM	0.64**	-0.30**	-0.06
JJA	0.54**	-0.35**	-0.29**
SON	0.75**	0.24**	0.30**

Table 6.3 – Correlation coefficients for changes in KZI, NWUIB T2 and global T2 future (projected) changes.

In Figure 6.13 we examine how the modulation of inter-annual variability in T2 by the KV may change in the future. If trends in the KZI and T2 series are not taken into account, there is generally a decrease in KZI/T2 correlation from the historical (grey) to future periods (orange). The starkest reduction in correlation occurs in summer. However, if the future series are detrended (blue), the correlations are essentially the same in the historical and future periods. This indicates that the modulating influence of the KV on inter-annual variability in T2 is projected to remain the same in the future, according to the CMIP ensemble. This is concordant with the finding that the GCMs provide a reasonable if slightly weak representation of connections between the KV and temperature for inter-annual variability, even if the strength of relationships between trends appears to be different.

Projection Spread

To account for variations in climate sensitivity between models, we identify the local temperature (T2) changes in the NWUIB associated with specific global changes, following the approach described in Section 6.3.1. Figure 6.14 shows that the residual spread in projected NWUIB T2 changes is large. While the ensemble minimum T2 changes are approximately equal to the global changes, most models project larger temperature increases. The degree of spread in the ensemble also increases substantially for larger global T2 change scenarios. Interestingly, some seasonality emerges whereby changes in winter and autumn exceed those in spring and summer.

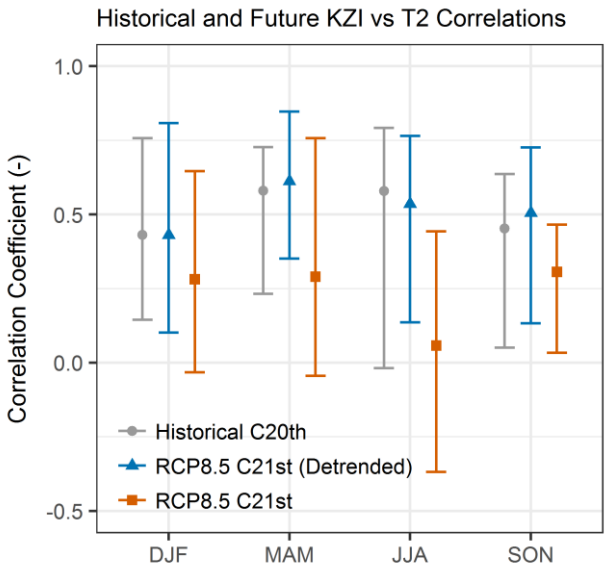


Figure 6.13 – Comparison of correlations between KZI and NWUIB T2 for historical (20th century) and future (21st century) under RCP8.5.

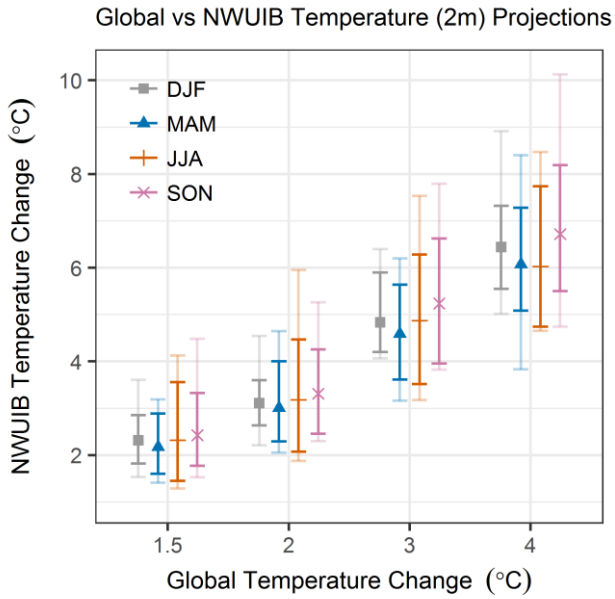


Figure 6.14 – Seasonal projected 2m temperature changes in the NWUIB for selected global temperature scenarios for the CMIP ensemble. The ensemble mean, 10-90th and total ranges are shown, the latter with transparency applied.

The question then arises as to what drives the residual spread in the ensemble, and whether the KV plays a leading role and how important other influences might be. To provide an

initial exploration of this, we used the ensemble to correlate a range of variables with seasonal T2 change factors for a 2°C warmer world. This tests whether a model with a high (low) value for a given variable tends to be associated with a relatively high (low) T2 change factor. We used KZS/KZI as a quantification of KV activity, as well as variables closely connected to the surface mass and energy balance, given their generally close relationship to climatological temperatures and variability. As well as KZS, the variables included are: precipitation, incoming shortwave radiation, albedo, incoming longwave radiation, sensible and latent heat fluxes, and temperature (T2). For each variable and season, the historical mean, standard deviation and correlation with T2 anomalies were calculated as potential covariates. We refer to the resulting combinations as variable/quantity pairs, e.g. precipitation/historical mean or albedo/historical standard deviation. The correlation between all combinations of these variable/quantity pairs were then calculated to give a correlation matrix.

In the following exploratory analysis, we focus primarily on those variable/quantity pairs significantly correlated with T2 change factors. As the correlation matrices show large numbers of significant correlations between various variable/quantity pairs, reflecting the interconnectedness and feedbacks of the climate system, this approach is substantially reductionist. However, it is intended primarily to facilitate identification of some critical issues. Further work could potentially explore dimensionality reduction and higher order relationships inter alia.

For winter, Table 6.4 shows that the variable/quantity pair most closely correlated with a model's T2 change factor is its incoming longwave radiation change factor. This is intuitive, given the typically strong relationship between surface incoming longwave radiation and the temperature profile in the lower troposphere, such that the two variables would be expected to show some coevolution. Interestingly, the next largest correlation is with the incoming shortwave radiation change factor, which is negatively associated with the magnitude of T2 change. One possible factor linking the longwave and shortwave radiation changes is cloud cover and so changes in cloud radiative effects. For example, models showing an increase in cloud cover would be likely to exhibit a decrease in incoming shortwave radiation but an increase in incoming longwave radiation. This is supported by the significant negative correlation between incoming shortwave and longwave change factors (-0.72) in the full correlation matrix. Thus, while high winter albedo ensures reflection of a high proportion of incoming shortwave radiation, incoming longwave radiation is readily absorbed and leads to surface and near-surface warming.

	Covariate	
Variable	**Quantity**	**Correlation Coefficient (-)**
Incoming Longwave Radiation	Change Factor	0.76
Incoming Shortwave Radiation	Change Factor	-0.55
Precipitation	Change Factor	0.54
Sensible Heat Flux	Historical Mean	-0.50

Table 6.4 – Winter (rank) correlations with NWUIB temperature (2m) change factor for a 2°C global warming scenario. Only correlations significant at the 95% level are shown.

From the full correlation matrix, we find that the incoming shortwave radiation change factor is significantly correlated with its historical mean (-0.74), standard deviation (0.58) and correlation with T2 variability (0.62), amongst other variable/quantity pairs. Thus, models with lower incoming shortwave radiation in the historical period, but higher variability and association with T2 anomalies, tend to show larger increases in shortwave radiation. If cloud is the relevant mechanism, this could be interpreted as less cloudy models in the historical period experiencing the biggest increases in cloud cover, while relatively cloudy models experience smaller increases (or potentially decreases) in cloud cover. It is therefore relevant in future work to consider how well the models correspond with available data products for incoming shortwave radiation and cloud cover, as a potential guide to model reliability.

Compared with winter, more variable/quantity pairs show significant relationships with T2 changes in summer (Table 6.5). Importantly, while not the largest correlation, the historical mean KZS of a model does show an association with its T2 change factor (Table 6.5), which was not the case for winter. The negative sign of the relationship suggests that models with relatively northerly SWJ position in summer exhibit lesser T2 increases, but the mechanisms could be complicated and involve covariation with multiple variables. However, notably larger correlation coefficients are found for the sensible heat flux, albedo and incoming shortwave radiation, especially in terms of their change factors. The full correlation matrix reveals that the sensible heat flux change factor is strongly negatively correlated with albedo change (-0.9). Due to the suppression of cold snow and ice surfaces on surface heating and turbulent exchange, it is intuitive that larger reductions in albedo permit higher sensible heat fluxes in summer, with a resultant positive feedback.

Variable	Covariate Quantity	Correlation Coefficient (-)
Sensible Heat Flux	Change Factor	0.84
Albedo	Change Factor	-0.83
Incoming Shortwave Radiation	Change Factor	-0.77
Albedo	Historical Mean	0.70
Sensible Heat Flux	Historical T2 Correlation	0.70
Albedo	Historical Standard Deviation	0.66
KZS	Historical Mean	-0.60
Precipitation	Historical Mean	-0.60
Incoming Longwave Radiation	Historical Mean	-0.59
Albedo	Historical T2 Correlation	-0.54
Temperature (2m)	Historical Standard Deviation	0.54
Incoming Longwave Radiation	Change Factor	0.54
Latent Heat Flux	Historical Mean	-0.51
Sensible Heat Flux	Historical Standard Deviation	0.51
Temperature (2m)	Historical Mean	-0.50

Table 6.5 – Summer (rank) correlations with NWUIB temperature (2m) change factor for a 2°C global warming scenario. Only correlations significant at the 95% level are shown.

The negative correlation of T2 changes with incoming shortwave radiation changes is interesting, as it suggests that larger reductions in incoming shortwave radiation (i.e. potential increasing cloudiness) are associated with larger T2 increases, which seems initially counterintuitive. This could be reflective of diurnal effects, such as increased cloud enhancement of longwave radiation at night, which could induce warming and potentially nocturnal melting, reduced albedo and thus more warming during the day. Alternatively, it could be an artefact of correlation or potentially reflect the influence of higher order associations. This could be explored in further work.

Given the clearer influence of albedo changes, which are correlated with its historical mean (-0.72) and standard deviation (-0.82), we now consider how it relates to proxy observations, leaving incoming shortwave radiation for further work. Figure 6.15 compares historical means and standard deviations of simulated albedo with albedo and temperature (2m, T2) change factors. Figure 6.15a shows that there is generally a non-linear relationship between historical albedo means and albedo change factors. T2 changes generally increase as albedo changes become more negative, with the largest albedo changes often found in models with the highest historical mean albedo. The range of estimates of historical mean albedo from MODIS remote sensing and ERA-Interim for the NWUIB differ notably, with most models falling somewhere in between. In most cases the estimates of T2 change factors based on RCP4.5 (circles) and RCP8.5 (triangles) are similar, suggesting that the approach to identifying change factors using model-specific time windows is sufficiently robust to natural variability.

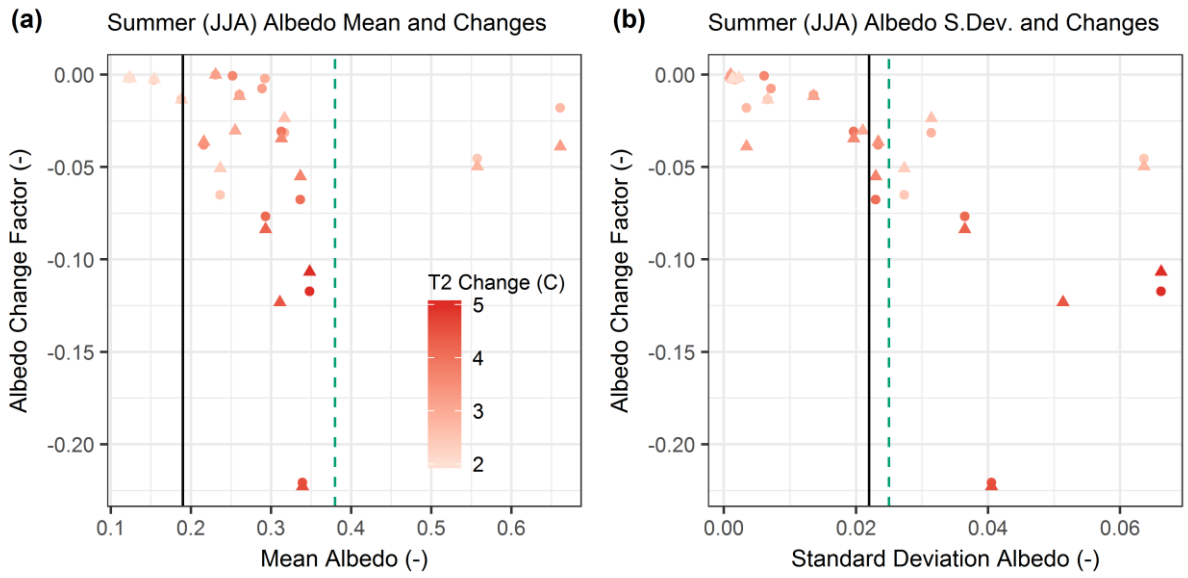


Figure 6.15 – Comparison of modelled mean (a) and standard deviation (b) of albedo with albedo change factor, as well as T2 change factor (colour). Black line denotes MODIS MCD43A3 reference and green dashed line shows ERAI reference.

Figure 6.15b shows a potentially more linear relationship between historical albedo standard deviation and projected albedo changes. Interestingly, the largest T2 changes are associated with some of the largest historical albedo standard deviations. Historical variability may therefore provide some guide to a model's likely projections. The difference in historical albedo standard deviation from ERA-Interim and MODIS is quite small, with most models falling outside the range this time. This offers potential to constrain the plausible range of summer T2 change factors. In particular, the analysis suggests that the largest changes may be

associated with deficient performance with respect to albedo representation, and probably snow albedo feedbacks, such that they may be considered less likely.

In order to understand the outliers in Figure 6.15, it is necessary to consider the co-variation of historical mean and standard deviation of albedo. Figure 6.16 plots these statistics against each other, along with T2 change factors. This shows that the models with very high historical mean albedo are associated with low T2 change factors, regardless of historical albedo standard deviation. It is likely that their albedos are so high that the sensitivity to changes in the future is limited compared with models showing lower albedos. For the other models, there is most likely a non-linear relationship between historical means and standard deviations of albedo. Models with low albedo tend to show low variability and low T2 changes, while models with higher means are associated with higher variability and higher T2 changes. The area contained by the MODIS (black line) and ERA-Interim (dashed green line) statistics in theory shows the models most consistent with observations, which appear to exhibit moderate temperature projections, as do the models surrounding this box. Potentially then this analysis shows that the more extreme low and high changes are less consistent with observations and physical process representation than the moderate T2 changes.

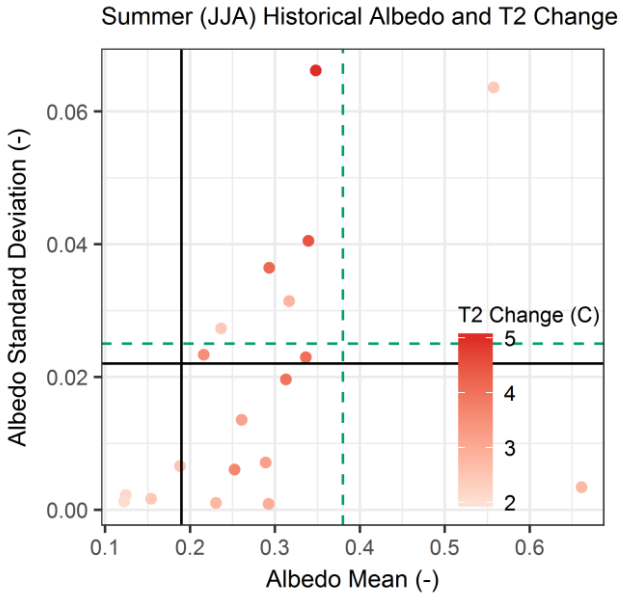


Figure 6.16 – As Figure 6.15 but showing historical mean and standard deviation of albedo together, with T2 change factors for RCP4.5 runs.

6.5 Discussion and Conclusions

The analysis above demonstrates that the KV system is generally present in the CMIP5 GCMs. The characteristic cycle of the KZS, an index for the position and intensity of the SWJ is well captured in most models, although a small southward/weak bias exists. This bias is reduced in the AMIP ensemble, suggesting that the small bias may be related to SST or sea ice biases in coupled models. However, the KZS bias is generally small and does not inhibit the virtually of the GCMS from capturing the basic structure of the KV. The correspondence of anomalous anticyclonic (cyclonic) circulation to positive (negative) KZI is present, as is the positive correlation with temperature through much of the troposphere. The magnitude of correlation does vary between models, as does the spatial extent of significant correlations. A summer contraction of the KV is also reasonably well simulated in the models, acknowledging inter-model variations in the centroid of the vortex and magnitude of correlation.

For the Karakoram, the GCMs generally show lower association between the KZI and near-surface air temperatures than reanalyses. However, the ensemble spread at this more local scale is reasonably large. There does not appear to be a simple connection between biases in KZS (or temperature) and the accuracy of correlations between KZI and temperature in the ensemble. Crucially, the GCMS do not fully capture the negative KZI trend of recent decades, but especially the summer temperature cooling present in observations. Therefore, while the models do reasonably in terms of representing the influence of the KV on inter-annual temperature variability, these results raise the question of why the historical trend in regional circulation and temperature does not appear to be adequately simulated. This has substantial implications, potentially reducing confidence in future projections if the models respond incorrectly to the changing conditions of the past decades.

In terms of projected changes, the GCM ensemble consistently shows warming in the Karakoram and UIB above the global average changes. After accounting for different rates of global temperature change, the results suggest that model skill for the historical period with respect to the KV plays some role in shaping the residual spread in the ensemble. Specifically, the mean KZS in summer in the historical period is significantly correlated with the degree of temperature change projected for the Karakoram in a 2°C warmer world. The magnitude of historical KZI/T2 correlations is not significantly correlated with projected temperature change in summer, while neither KZS nor KZI/T2 correlation shows notable association with

temperature change in winter. Moreover, it appears that inter-model variation in local influences on the SEB have a larger influence. In winter, there is evidence to suggest that differences in projected cloud cover is important, although the mean partitioning of sensible and latent heat fluxes in the historical period is also crucial. In summer, feedbacks between incoming shortwave radiation, albedo and sensible heat fluxes are critical.

The results provide some suggestion that temperature change projections can be constrained with reference to representation of some historical land surface states and processes, particularly related to surface albedo. This is essentially in line with findings on the role of snow albedo feedbacks at larger scales, in terms of how inter-model variation offers the potential for applying observational constraints to projected changes (Hall and Qu 2006; Essery 2013; Qu and Hall 2014). This analysis suggests that similar methods may be applicable at the regional scale. In contrast to approaches taking cold/warm and dry/wet extremes to produce a very wide envelope of climate projections (e.g. Lutz et al. 2016a,b), there may thus be scope to reduce the range of possible climatic and hydrological futures projected for the Karakoram, UIB and neighbouring basins draining the Himalaya and Tibetan Plateau.

Chapter 7

Climate and Hydrological Projections for Global Warming Targets

7.1 Introduction

The coming years will be a critical period for devising and implementing climate change mitigation policy to meet the aims of the Paris Agreement. This is particularly so given the short timescales for action (Figueres et al. 2017), currently insufficient national emissions pledges (Hulme 2016), levels of committed warming from past emissions (Mauritsen and Pincus 2017) and uncertainty on whether the large role typically assumed for negative-emissions technologies is realistic (Anderson and Peters 2016). For mitigation policy and action to progress in the face of these barriers, analyses are required to inform policy-makers of the possible differences in impacts between climate stabilisation scenarios. It is therefore a problem that the relative risks and opportunities of climate stabilisation at the global temperature targets referred to in the Paris Agreement – 1.5 and 2°C above pre-industrial levels – have received relatively little attention until recently (Mitchell et al. 2016). The Intergovernmental Panel on Climate Change (IPCC) report on impacts of 1.5°C warming sought to address this gap (IPCC 2018), but there remains significant scope for more detailed consideration of regional impacts for the Himalaya and Tibetan Plateau.

For the focus of this study, the Upper Indus Basin (UIB), previous analyses of climate change impacts on basin hydrology have largely focused on fixed time slices for different emissions trajectories, rather than climate stabilisation targets. For example, several studies have suggested that summer river flows in the basin may initially increase over the coming decades under Representative Concentration Pathway (RCP) RCP4.5 and RCP8.5 pathways, due to higher glacial melt in a warming climate (e.g. Immerzeel et al. 2013; Lutz et al. 2014; Soncini et al. 2014; Ali et al. 2015). These projections also generally posit that glacier melt will subsequently decrease due to loss of glacier mass, but that this may be at least partly compensated for in total river flows by summer precipitation increases (Palazzi et al. 2015). The studies also suggest that the onset of seasonal snowpack melting may shift earlier in the year. However, Lutz et al. (2016a) emphasised the very high uncertainty in such projections, largely due to climate model spread. The apparent discrepancy between recent observed changes in the UIB and the warming predicted by climate models over both historical and

future periods has also been noted as a possible reflection of limits to climate model fidelity in this region (Hasson 2016b).

In one of the few studies evaluating the possible impacts of climate stabilisation targets in the region, Kraaijenbrink et al. (2017) modelled glacier mass trajectories across the whole of the Himalayan arc. Based on a small subset of Coupled Model Inter-Comparison Project Phase 5 (CMIP5) (Taylor et al. 2012) model outputs, their results suggested that, relative to present conditions, central estimates of glacier mass loss under a 1.5°C global warming scenario could be around 17% in the Karakoram, 48% in the Hindu Kush and 42% in the western Himalaya by the end of the century. This is in contrast with a projected region-wide mass loss of 36%, which increases to around 51% and 68% by the end of the century in simulations using RCP4.5 and RCP8.5 pathways, respectively. As Kraaijenbrink et al. noted, substantial uncertainty arises due to the region's patchy and uncertain mass balance observations, against which the model was calibrated. There could also be notable implications from inaccuracy in the initial ice thicknesses prescribed, which can differ greatly as a function of alternative estimation methods (Farinotti et al. 2017).

While Kraaijenbrink et al.'s (2017) study usefully highlights the possible differences in glacier response across the region, as well as the substantially lower mass loss under a 1.5°C global warming scenario, the hydrological dimensions of these changes have not been explored. Therefore, we explore here the impacts of selected climate stabilisation scenarios on the hydrology of the UIB. There are many uncertainties in this type of modelling exercise, deriving from climate models, downscaling methods and hydrological models (Clark et al. 2016), such that it quickly becomes intractable to evaluate all of the possible combinations of methods and scenarios quantitatively. With this in mind, and in combination with the fact that climate stabilisation studies for UIB hydrology have not been undertaken previously, we start fairly simply in terms of methods employed and hydrological responses investigated. The intention is thus to provide an initial scoping exercise, against which the results from more sophisticated analyses of projections and uncertainties could be compared in further work.

There are three components to the analysis presented here. Firstly, we assess the climate model runs conducted under the auspices of the Half a degree Additional warming, Prognosis and Projected Impacts (HAPPI) protocol (Mitchell et al. 2017). In contrast to CMIP5, the HAPPI simulations have not been evaluated in the UIB region before, such that we appraise their basic suitability for application here. Secondly, we analyse the climate stabilisation

projections for the UIB from HAPPI, comparing them with outputs from CMIP5. Finally, we explore the hydrological implications of different climate stabilisation scenarios. The emphasis here is on some basic water resources considerations, namely changes in annual flows and shifts in seasonality. Given its reasonably good performance in Chapter 5, the results focus on simulations with TOPKAPI-ETH.

7.2 Data and Methods

7.2.1 HAPPI Evaluation

HAPPI Background

The HAPPI protocol was designed to overcome some of the difficulties in delineating the impacts of 1.5 and 2°C warming scenarios from other climate model outputs (Mitchell et al. 2017). These challenges include separation or identification of any signal associated with an additional half degree warming from the uncertainty and internal variability in coupled atmosphere-ocean simulations under comparatively low emissions trajectories. HAPPI attempts to address this issue with relatively large ensembles of integrations using perturbed initial conditions in atmosphere-only models. In this study, we focus on the Tier 1 experiment in HAPPI. Happily, the trajectories in the CMIP5 ensemble stabilise at around 1.5°C relative to preindustrial conditions in RCP2.6, such that the simulations in the Tier 1 experiment for the 1.5°C scenario use the end-of-century anthropogenic radiative forcing from RCP2.6. With no similarly convenient option for the 2°C scenario, Mitchell et al. (2017) undertook radiative forcing from RCP2.6 and RCP4.5 weighted by their respective global temperature changes at stabilisation. The Tier 1 experiment in HAPPI uses single estimates of sea surface temperatures (SST) and sea ice concentrations by perturbing historical observations using CMIP5 outputs, as described in Mitchell et al. (2017). The models used in the Tier 1 experiments are summarised in Table 7.1.

The HAPPI simulation outputs have been used in a growing number of applications. These include its extensive use in the recent IPCC report on 1.5°C warming (IPCC 2018). Of particular relevance to the broader region surrounding the UIB, HAPPI applications (hereafter HAPPlications) include an assessment highlighting South Asia as a possible hotspot for increasing extremes in river flows (Paltan et al. 2018). Supporting these findings, Lee et al. (2018) also noted larger increases in mean and extreme precipitation during the South Asian monsoon under a 2°C warming scenario in both HAPPI and CMIP5. Indeed, a strong part of

the rationale for the large-ensemble approach in HAPPI was to facilitate such evaluations of changes in regional extremes by concatenating runs to produce sufficiently long time series (Mitchell et al. 2017). However, Fischer et al. (2018) demonstrated that using prescribed SSTs, as in HAPPI, can lead to projected changes in extremes that are biased high and subject to a narrower range of uncertainty than in equivalent analyses coupled atmosphere-ocean models. Recognising these possible issues with prescribed SSTs, Mitchell et al. (2017) outlined Tier 2 experiments to evaluate SST-related uncertainties. However, these simulations were not completed in time for inclusion in this work unfortunately. As such, and for comparison with previous studies, we also consider CMIP5 outputs in our projections, as described below.

Model	Horizontal Resolution	Runs per Scenario	References
CAM4	2 x 2°	500	Neale et al. (2013)
CAM5.1.2	0.25 x 0.25°	5	Wehner et al. (2014)
CanAM4	T63	100	von Salzen et al. (2013)
ECHAM6.3-LR	T63	100	Stevens et al. (2013)
MIROC5	1.5 x 1.5°	100	Watanabe et al. (2010)
NorESM1-HAPPI	1.25 x 0.94°	125	Bentsen et al. (2013)

Table 7.1 – Summary of GCMs in the HAPPI Tier 1 experiment used in this study.

HAPPI Evaluation

For the HAPPI model evaluation, we focus on basic precipitation and temperature climatology. The evaluation of simulated climatology in the HAPPI models uses local climate observations from 13 stations covering a range of elevations in the north-west UIB (NWUIB hereafter), as introduced in Chapter 3. The purpose of this is to provide confirmation that the models simulate the UIB climate to at least a similar degree of accuracy as the CMIP5 models. This may be expected, given that the models used in HAPPI are (or are closely related to) the atmospheric components of CMIP5 models (Mitchell et al. 2017). The main exception to this is the high resolution (0.25°) version of the Community Atmosphere Model v5 (CAM5.1.2) global climate model (GCM). This model thus attains a substantially higher spatial resolution than even the Coordinated Regional Climate Downscaling Experiment (CORDEX) regional climate model (RCM) simulations for South Asia. As such, we

performed a more detailed evaluation of CAM5.1.2 than the other, coarser models in the HAPPI ensemble. For this we transferred some of the evaluation approaches used for the High Asia Refined Analysis (HAR) (Maussion et al. 2014) in Chapter 3. This evaluation is provided in Appendix D (Section D.2). The main finding from the evaluation of the CAM5.1.2 outputs was that the advantages of its higher resolution may be at least partly counteracted by issues associated with the direct transfer of parameterisation tuning from coarser versions of the model (Wehner et al. 2014). As such, it is unclear that the model adds particularly to the HAPPI ensemble at present, while some aspects of model behaviour hint at possible issues in the simulations. Therefore, we do not use the CAM5.1.2 outputs as a basis for projections in this chapter.

7.2.2 Climate Scenarios

CMIP5

While the CMIP5 ensemble was not specifically designed to answer questions about 1.5 and 2°C global warming scenarios, it provides a useful complement to HAPPI (see discussion above), hence the inclusion of both in the recent IPCC report on 1.5°C warming (IPCC 2018). However, this raises the question of how best to use the CMIP5 outputs to derive climate scenarios. This has received recent attention in the literature, such that we adopt what Tebaldi and Knutti (2018) term the "time-shift" approach. The method involves identifying a time window in the transient CMIP5 simulations when the global mean temperature is at the desired target relative to pre-industrial levels. This time window can then be compared to a historical baseline period.

The method has been applied in an increasing number of studies (Fischer and Knutti 2015; Schleussner et al. 2016; King et al. 2017), but one possible drawback is that different models reach the specified global temperature targets at different points and potentially in response to different forcing. This could partly reflect differential localised aerosol influences or natural variability at multi-decadal time scales (Mitchell et al. 2017; James et al. 2017). However, taking advantage of bespoke simulations with a fully coupled atmosphere-ocean model, Tebaldi and Knutti (2018) recently demonstrated that simple pattern scaling or time-shift approaches are likely sufficient for many applications. Specifically, they found that the errors from these methods are small in comparison to the range of forced response in a multi-model ensemble, as well as the internal variability overlying forced response in a given model. This may apply particularly for patterns of temperature change.

To identify the required time periods from each CMIP5 model (see Chapter 6), we used a 30-year moving average with time series of simulated global mean temperatures. Outputs from both RCP4.5 and RCP8.5 runs were used to provide some elementary consideration of the differences in regional changes associated with different runs of a given model. Ideally we would consider multiple runs with perturbed initial conditions for each model/RCP combination, but not all models appear to have sufficient simulations available. However, we did additionally identify the time windows corresponding with 3 and 4°C global temperature change scenarios. This provides some means of comparing the ambitious 1.5 and 2°C targets referenced in the Paris Agreement with outcomes perhaps more in line with current trajectories.

Climate Scenario Downscaling

As this study represents an initial exploration of the hydrological implications of different climate stabilisation scenarios, we adopt a simple approach to developing the climate input fields required to drive hydrological models. Specifically, we use a delta change approach to perturb historical input fields based on the differences in monthly climatology between historical baseline and target scenarios. Change factors were calculated based on spatially averaging for grid cells overlapping the north-west UIB (hereafter NWUIB) domain introduced in previous chapters. The historical input fields are derived from the HAR, following the methodology discussed in Chapters 4 and 5. The delta change method has been very widely applied in the Himalayan region and elsewhere, being simple to implement and interpret. However, it does not capture differential changes across the (joint) distribution(s) of pertinent climate fields. As such, the method does not account for possible alterations to variability, extreme events or temporal patterns in future climates, such as changes to the duration or frequency of wet and dry spells (Fowler et al. 2007; Maraun et al. 2010). As applied here, the method also ignores spatial (vertical) variation in projected changes at sub-NWUIB scales. This could be a subject of further work when adequate high resolution simulations are available.

Some of these issues are partially mitigated by the focus of this study. Rather than examining changes in extremes, changes in overall flow volumes and seasonality are foregrounded. The temporal offset of the primary precipitation (winter/spring snowfall) and runoff (summer) seasons in the UIB also reduces the sensitivity of the system response to possible changes in the temporal distribution of precipitation at a sub-seasonal level during winter and spring.

Clearly though the joint distribution of precipitation and temperature remains important for possible phase changes. Moreover, preliminary analysis of the HAPPI ensemble suggested that, for a given month, change factors are fairly similar when calculated for different percentiles of the distribution arising from inter-annual variability, with differences in response across the distribution generally small relative to the mean change (Appendix D, Section D.1). Nevertheless, we emphasise that the approach adopted constitutes an initial analysis. Future work could compare the results here with more sophisticated downscaling approaches, such as a weather generator (Kilsby et al. 2007; Forsythe et al. 2014) or recently developed quasi-dynamical downscaling models that may be suitable for mountainous terrain (Gutmann et al. 2016).

Experiment	Climate Model Ensemble	Global Temperature Targets (°C)	Variables Perturbed	Change Factor Calculation
1	HAPPI	1.5, 2	Temperature	Means of initial conditions ensemble for each model/scenario
2	HAPPI	1.5, 2	Temperature, precipitation	Means of initial conditions ensemble for each model/scenario
3	HAPPI	1.5, 2	Temperature, precipitation	100 randomly sampled combinations of runs from initial conditions ensemble for each model/scenario
4	CMIP5	1.5, 2	Temperature	Separate calculations for RCP4.5 and RCP8.5
5	CMIP5	1.5, 2	Temperature, precipitation	Separate calculations for RCP4.5 and RCP8.5
6	CMIP5	3, 4	Temperature	Separate calculations for RCP4.5 and RCP8.5
7	CMIP5	3, 4	Temperature, precipitation	Separate calculations for RCP4.5 and RCP8.5

Table 7.2 – Summary of climate scenarios used for hydrological projections.

Moreover, while we are not using a stochastic downscaling approach, such as a weather generator, we are able to make use of HAPPI's large ensembles to take uncertainty into account to a degree. This is achieved by randomly sampling runs from the initial conditions

ensembles for each scenario, in order to calculate a range of change factors (CFs) with the sampled combinations. With less runs available for CMIP5, both RCP4.5 and RCP8.5 outputs were used to calculate CFs for each model, thereby providing a very elementary indication of uncertainty that could be developed further in future work.

These strategies lead to the ensembles of climate scenarios described in Table 7.2. These are the scenarios investigated in the hydrological projections in this chapter. Notably they include scenarios perturbing both precipitation and temperature, as well as simulations involving only temperature changes. The purpose of this is to delineate the role of the temperature forcing, given the substantial variability associated with precipitation projections between models and indeed between different runs of any given model. Based on the analysis of HAPPI climate scenarios in Section 7.3.2, we focus solely on precipitation and temperature. CMIP5 outputs are used to provide 3 and 4°C global warming scenarios for comparison.

While the results in Chapter 6 start to identify possible routes to constraining climate model ensembles based on process fidelity in this region, we retain the array of models in the HAPPI and CMIP5 ensembles in development of the climate scenarios used in this chapter. In part this is because further work is required on observational constraints. Examining the range of possible climate changes also represents a precautionary, conservative approach. This may be valuable in this application, given that it is one of the first attempts to focus on climate stabilisation scenarios in the UIB.

7.2.3 Hydrological Projections

Following the baseline setup described in Chapter 5, the TOPKAPI-ETH model was used to provide hydrological projections for each of the experiments detailed in Table 7.2. TOPKAPI-ETH was shown to provide reasonable performance for both snow-dominated and heavily glaciated UIB sub-basins in Chapter 5. Moreover, the baseline parameter choices for the enhanced temperature index (ETI) melt model appear to be reasonable in terms of temperature sensitivity. This is particularly so compared with either a simple temperature index (TI) model or possible alternative choices for the ETI model. The model configuration applied here therefore largely follows that described in Chapter 5. As per Chapter 5, we simulated both the snow-dominated Astore and heavily glaciated Hunza sub-basins. As such, we leave it for further work to compare TOPKAPI-ETH projections with those made using models that solve the surface energy balance. Ideally, a weather generator or quasi-dynamical downscaling approach could be used to provide climate inputs in such a study. This would help to account

for the spatial, temporal and inter-variable dependence structures of the multiple climate fields required.

As the focus of this study is climate stabilisation targets, we cycled the climate input time series multiple times to ensure that dynamic equilibrium conditions were reached. With glaciers representing the major storage term in the UIB, neutral glacier mass balance at the catchment-scale was taken to indicate dynamic equilibrium. The analysis is based on the last cycle of climate inputs for any given model run. The climate inputs cycled are for the period 1961-2010, using the method described in Chapter 5 to extend the HAR time series based on anomaly time series from long-term in-situ observations.

7.3 Results and Discussion

This section begins with an evaluation of the model outputs in the HAPPI ensemble. This begins with the CAM5.1.2 model, which is analysed separately owing to its substantially higher spatial resolution, as noted in Section 7.2.1. After then assessing the climatology in the coarser models, the climate projections associated with the selected climate stabilisation targets are characterised. Finally, the hydrological implications of these climate scenarios are explored.

7.3.1 HAPPI Evaluation

Figure 7.1 compares the annual cycle of precipitation in the HAPPI models with station observations. This confirms that the models capture the winter/spring peak in NWUIB precipitation, albeit to varying degrees. CAM4 exhibits a particularly flat annual cycle, which indicates some possible issues capturing the balance between westerly disturbances and possible monsoon incursions. The higher resolution CAM5.1.2 model does show slightly more agreement with observations than the lower resolution CAM4, although the differences between these models go beyond resolution. Some of the models show a relatively large fraction of annual precipitation occurring in autumn compared with observations, while there is some variation in the extent to which the models simulate a secondary peak in summer. As such, the HAPPI models likely show a comparable level of performance in simulating precipitation seasonality as the CMIP5 ensemble. Previous CMIP5 analyses show that the models in CMIP5 are generally able to reproduce the differentiation of westerly- and monsoon-dominated precipitation regimes with progression eastwards along the Himalayan arc (e.g. Palazzi et al. 2015). However, a wet bias might be evident, along with some under-

estimation of seasonal variation that may also be present in the HAPPI models, but the observational reference datasets used in these evaluations are not without limitations (Palazzi et al. 2015; Hasson et al. 2018).

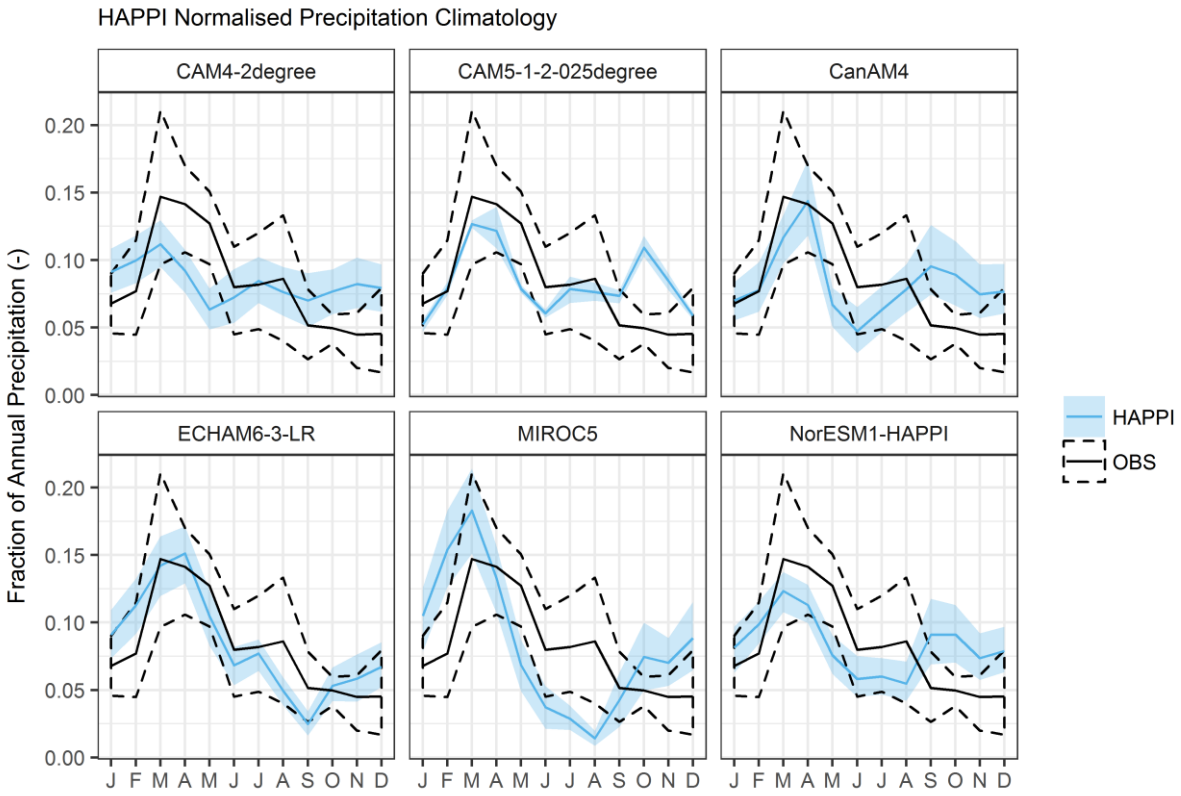


Figure 7.1 – Comparison of normalised annual cycle of precipitation for HAPPI and station observations. Shading corresponds with the 10-90th percentile range for the HAPPI initial conditions ensemble, while the range for observations corresponds with the station ensemble. Solid lines show the mean.

Figure 7.2 provides a similar comparison but this time for near-surface air temperature. This illustrates that the models all reproduce the annual cycle in general terms. The models generally show a larger amplitude cycle than observations, which may in part stem from the mismatch in scales between point measurements and coarse climate model grid cells. The modelled temperature often continues to rise above observations in late spring and early summer. This is particularly pronounced in MIROC5, which shows an especially steep peak. ECHAM6.3 shows greatest agreement with the observed normalised temperature cycle. Overall, however, this level of performance is similar to that reported for the CMIP5 ensemble, with the models generally subject to a cold bias over the Himalaya and Tibetan Plateau (Hasson et al. 2018). Given that the HAPPI models show reasonable performance

comparable with CMIP5, we consider them acceptable as one source of data for climate and hydrological projections in this chapter.

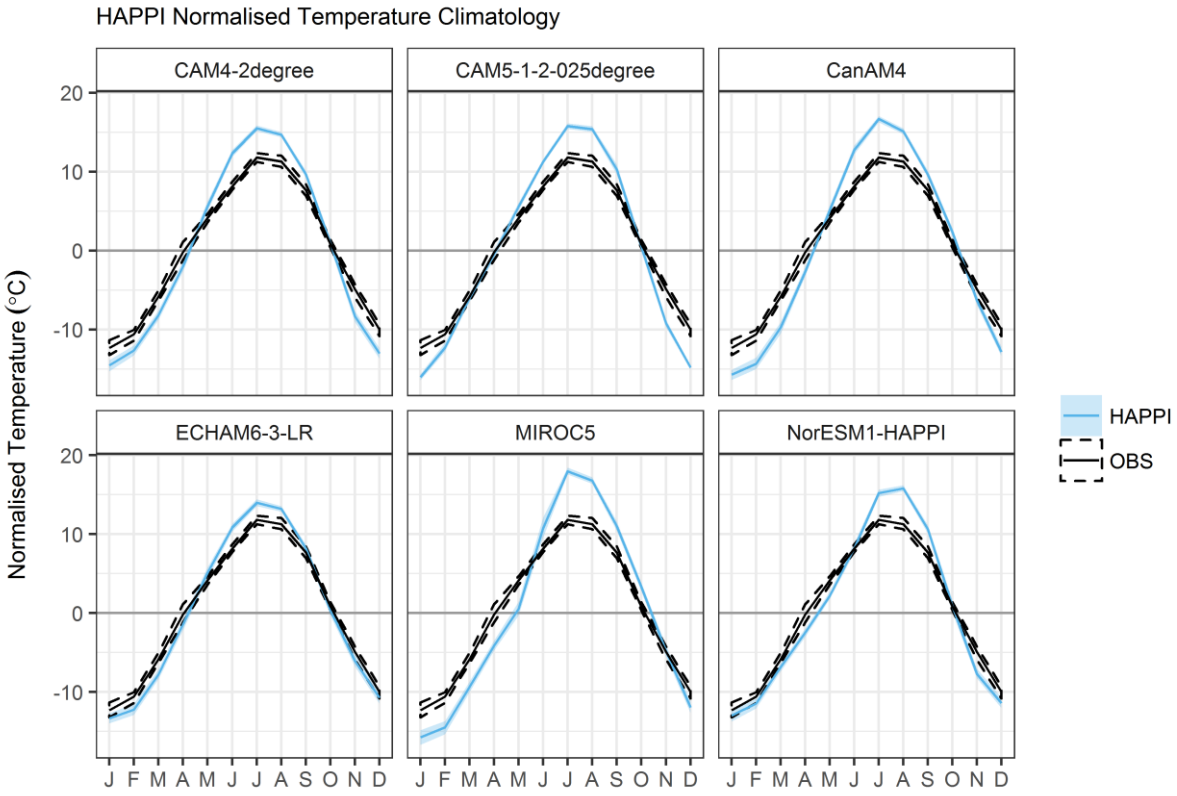


Figure 7.2 – As Figure 7.1 but for near-surface air temperature, with normalisation based on subtraction of the annual mean.

7.3.2 Climate Projections

HAPPI – Mean Changes

The analysis of climate projections for the UIB under 1.5 and 2°C global warming scenarios begins with change factors for the monthly climatology for HAPPI models. Figure 7.3 and Figure 7.4 display the change factors associated with each model for the 1.5 and 2°C scenarios, respectively. Expressed as heat maps, the change factors are calculated for each month by taking means using all runs in the initial conditions ensembles for each model/scenario combination separately. For temperature, this shows that increases are projected for both minimum (night-time) and maximum (day-time) temperatures in both scenarios for virtually all model/month combinations. This seems an intuitive outcome for warmer worlds, but we note the findings from Chapter 6 and Hasson (2016b) that such model behaviour may represent a departure from historical trends for the summer months.

Interestingly, the lowest warming in the 1.5°C scenario is associated with the NorESM1 model in the summer months. This would seem to fit with the summer cooling/stable temperatures observed in recent decades, linked with the Karakoram Vortex (KV – see Chapter 6), apart from the fact that the NorESM1 model appears to exhibit relatively poor performance in simulating the relationship between the KV and near-surface air temperature, as demonstrated in Chapter 6.

Comparing Figure 7.3 and Figure 7.4 suggests that, intuitively, larger projected temperature changes are found in the 2°C scenario. The differences between the scenarios average 0.7 to 0.8°C for minimum and maximum temperatures, respectively, with a range of differences from around 0.5 to 1.4°C. As such, the temperature differences between the 1.5 and 2°C global warming scenarios exceed the 0.5°C global difference for virtually all combinations of models and months. There is thus regional amplification of the global warming signal afflicting the UIB in these projections.

Figure 7.3 and Figure 7.4 do not appear to show very clearly defined patterns of seasonal variation in simulated temperature changes across temperature changes. The highest temperature changes are associated with the MIROC5 model, which reach 4°C in early summer in the 2°C global warming scenario. Note that this change is expressed relative to the recent historical baseline temperatures of the HAPPI runs, not the approximately 1°C cooler pre-industrial conditions, making this temperature change even more striking. While signals across the HAPPI models may not be clear, the patterns of seasonal variation are similar for any given model in both of the global warming scenarios. The 2°C scenario thus has the effect of largely amplifying the intra-annual cycle of change factors in the 1.5°C scenario. For some models, the patterns of seasonal variation are similar for minimum and maximum temperatures (e.g. NorESM1, ECHAM6-3-LR and CAM4), whereas for the others there is a suggestion of diurnal asymmetry in response in some months.

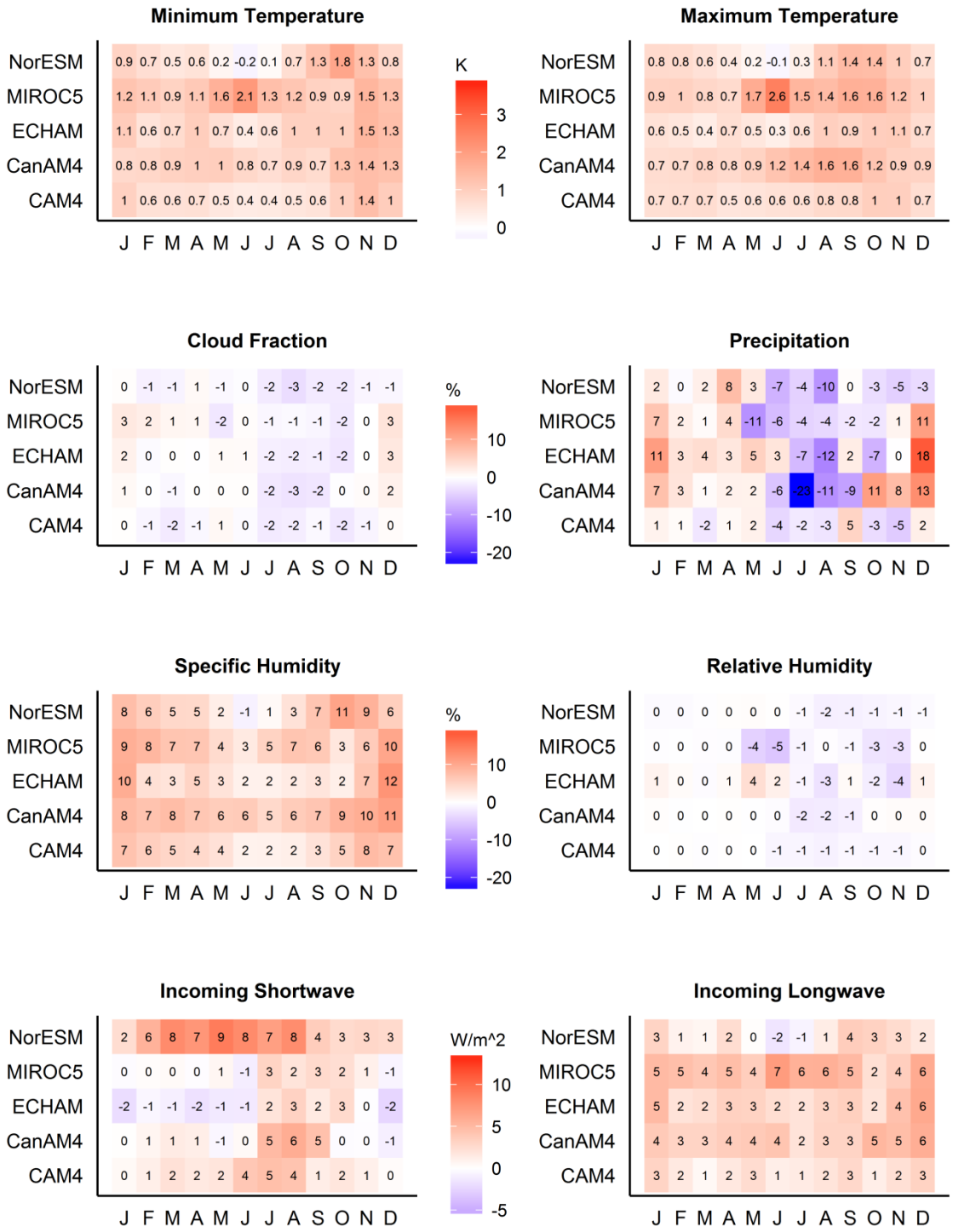


Figure 7.3 – Monthly change factors for HAPPI model initial conditions ensemble means for the 1.5°C warming scenario. Change factors are calculated relative to the recent historical baseline, which is approximately 1°C warmer than pre-industrial conditions.

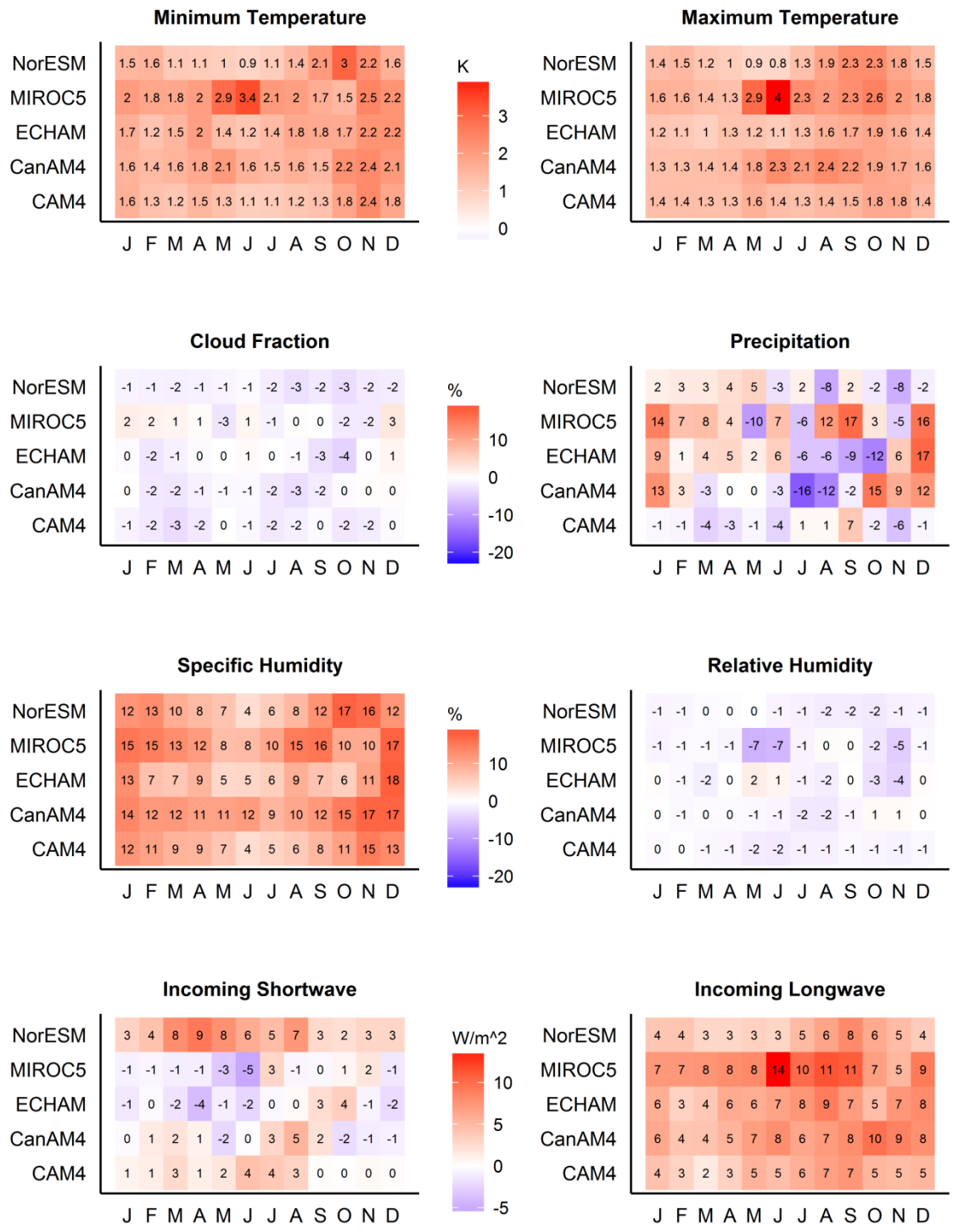


Figure 7.4 – As Figure 7.3 but for the 2.0°C warming scenario.

Compared with temperature, precipitation shows more mixed changes in the future scenarios in Figure 7.3 and Figure 7.4, with both positive and negative shifts being projected in different months in different models. These projected changes range from around -20 to +20% in some model/month combinations, but in most months changes are less than approximately

+/-10%. There is some suggestion of lower precipitation in the 1.5 and 2°C scenarios relative to the baseline in the summer and early autumn, although this diminishes a little in the warmer of the two cases. There is also an indication of some possible precipitation increases in the winter months. However, it seems unclear how significant these signals may be from examining changes based on means of initial conditions ensembles alone. This point is addressed shortly.

Several additional variables relevant to the surface energy balance are also presented in Figure 7.3 and Figure 7.4. For example, the figures confirm that specific humidity consistently increases in the warmer climates, while relative humidity remains largely unchanged for most months in most models. The main exception to this is the interesting reduction in relative humidity accompanying the large temperature increases in May and June in MIROC5. There is some correspondence between the magnitude of temperature and specific humidity changes in general, as might be expected. Possible implications for hydrological modelling applying such change factors would include latent heat flux calculations by altering near-surface humidity gradients, although the implications would potentially vary as a function of surface type and time of year.

Figure 7.3 and Figure 7.4 also indicate that projected changes in simulated mean cloud cover in both of the future scenarios are relatively small in the HAPPI models. For most months and models, the changes are negative if present at all, especially in the 2°C scenario. There is generally some negative association between changes in cloud cover and changes in incoming shortwave radiation, with increased cloud cover leading to lower shortwave radiation received at the surface and vice versa. However, the relationship between the magnitude of cloud changes and the magnitude of shortwave radiation changes may not be straight-forward. NorESM1 simulates the largest increases in incoming shortwave radiation, which are counteracted by its comparatively limited changes in incoming longwave radiation. Otherwise, the models tend to simulate incoming longwave radiation increases that exceed any shortwave changes, particularly in the 2°C scenario. The annual cycles of variation in incoming longwave radiation changes bear some resemblance to the patterns in temperature and specific humidity, but again the relationships are not just a simple mapping.

From the perspective of hydrological modelling implications, this raises the point that accounting for incoming longwave radiation changes could be useful in simulations of the full surface energy balance. Taking snow as an example, winter increases in incoming longwave

radiation might be largely counteracted by higher outgoing radiation from higher surface temperatures. However, the sensitivity to incoming longwave radiation may increase with progression into spring and summer, as temperatures and incoming shortwave radiation rise while the snow surface temperature is constrained to 0°C. The ideal route to investigating these issues further would be with a weather generator, potentially conditioned on high resolution coupled land-atmosphere modelling rather than sparse observations, or dynamical or quasi-dynamical downscaling. These approaches could provide more realistic, physically consistent fields than simple change factor approaches.

HAPPI – Initial Conditions Ensemble Changes

While Figure 7.3 and Figure 7.4 characterise changes associated with the means of the initial conditions ensembles in the HAPPI project, it is also instructive to make use of the large ensembles to consider the distributions arising from internal variability. As such, we now examine distributions of climatological monthly means in the historical and warming scenarios for the critical input variables for TOPKAPI-ETH, namely precipitation and temperature. To reiterate, this is not a consideration of inter-annual variability, but in effect an assessment of the uncertainty in climatological values arising from internal variability induced by perturbations of initial conditions. The aim is to provide some insight into the robustness of the mean changes apparent in Figure 7.3 and Figure 7.4.

Beginning with precipitation, Figure 7.5 expresses the distributions of climatological monthly means from the initial conditions ensembles as anomalies relative to the historical scenario mean. This suggests that the overall differences between the monthly mean precipitation in the historical and future scenarios in HAPPI are fairly limited. Most of the distributions in Figure 7.5 are very similar to each other, typically peaking around the 0% anomaly relative to the historical mean. The substantial overlap between distributions raises the question of whether any clear shifts in climatological precipitation in warmer worlds are apparent, once the influence of internal variability is taken into account. Interestingly, the CAM4 model shows particularly striking similarity between the scenarios. CAM4 has 500 runs for each scenario in its ensemble, compared with 100 runs for the other models, such that its greater coherence could be at least partly a function of having a larger sample with which to characterise the distribution. While the other models show a range of precipitation changes from -23 to +18% (looking at all models and months), CAM4 precipitation change factors in Figure 7.3 and Figure 7.4 are consistently lower (range of -6 to +7%). In contrast, the most

discernible differences between scenarios occur for CanAM4 in summer, although the other models do not show corroborating patterns.

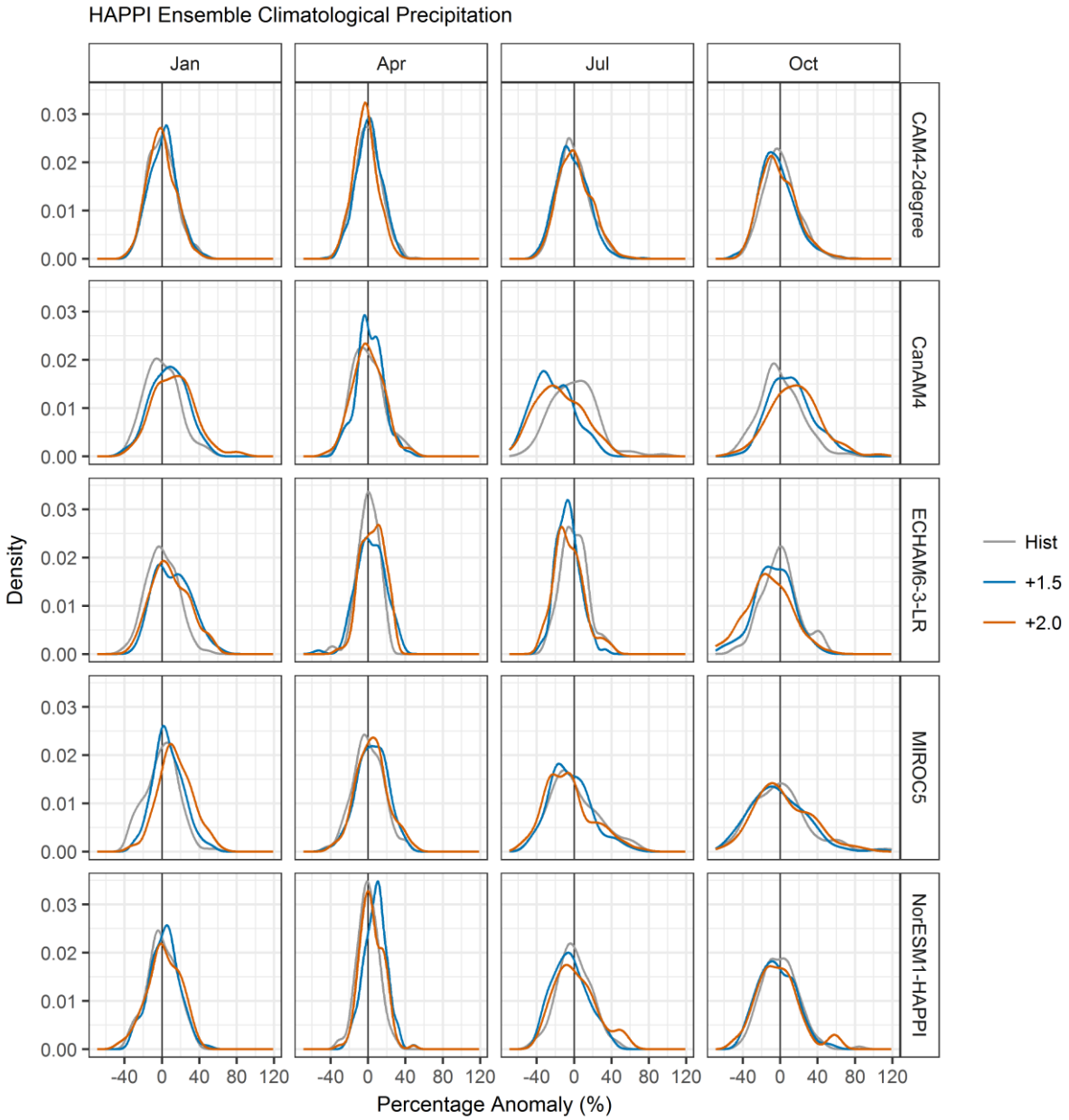


Figure 7.5 – Distributions of climatological monthly mean precipitation based on HAPPI model (initial conditions) ensembles for selected months. Distributions are estimated from the climatological means associated with each run for a given model/scenario combination.

Clearer differentiation between scenarios is evident in Figure 7.6 and Figure 7.7 for distributions of monthly means of daily minimum and daily maximum temperatures, respectively. This applies to almost all month/model combinations shown, supporting the findings from Figure 7.3 and Figure 7.4 that robust shifts in temperature changes are detectable in the HAPPI ensemble. Minimum temperatures represent an interesting case, as

there is notable difference in the shape of the distributions between the winter/spring and summer example months (Figure 7.6). While the distributions for winter/spring are comparatively flat, the spread reduces very markedly in the summer months. This implies that perturbing initial conditions can have a particularly large effect on simulated climatological minimum temperatures in winter and spring. In contrast, climatological minimum temperatures in summer show much lower sensitivity to initial conditions perturbations and thus internal variability.

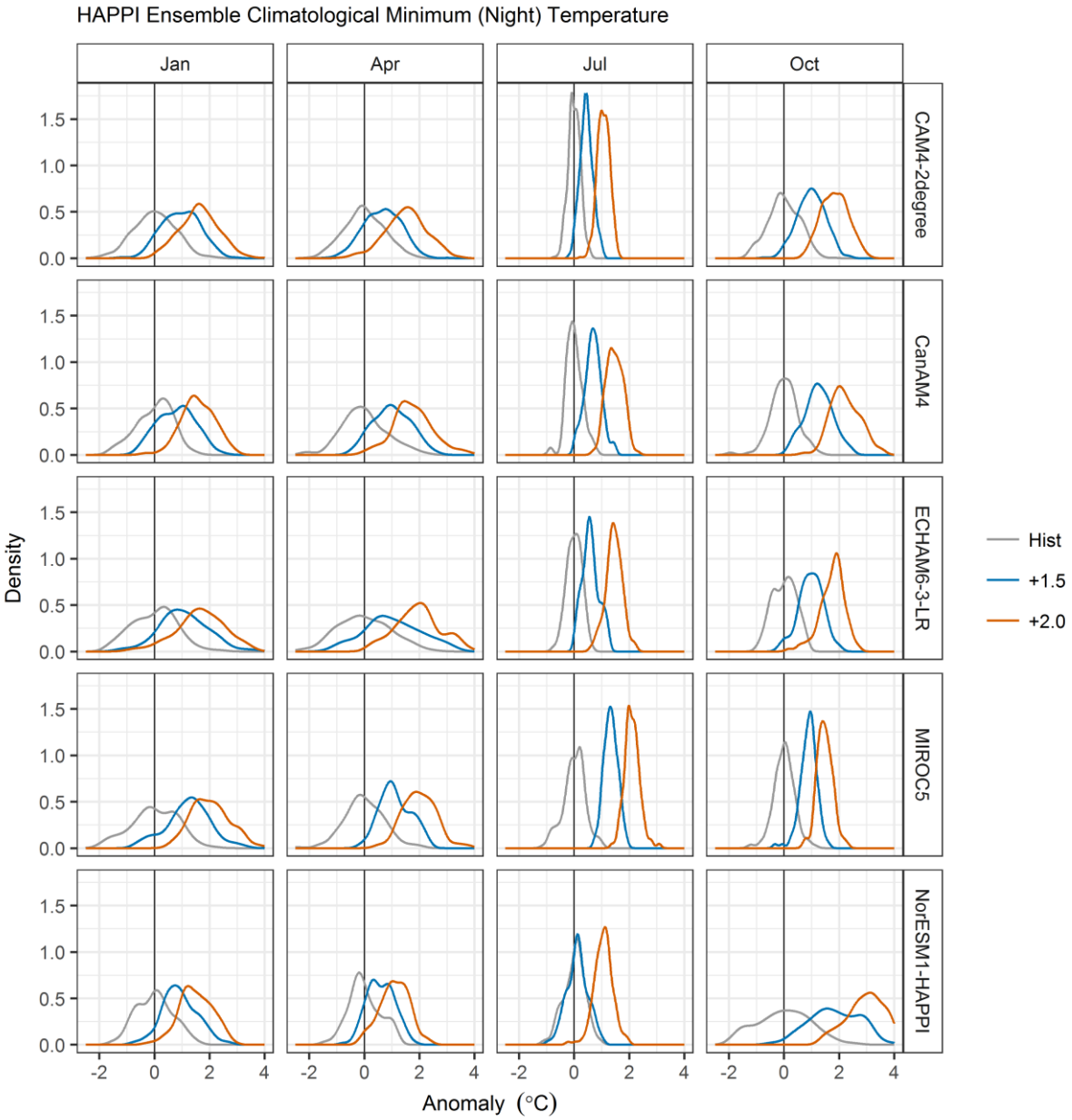


Figure 7.6 – As Figure 7.5, but for climatological monthly means of daily minimum (night-time) temperatures.

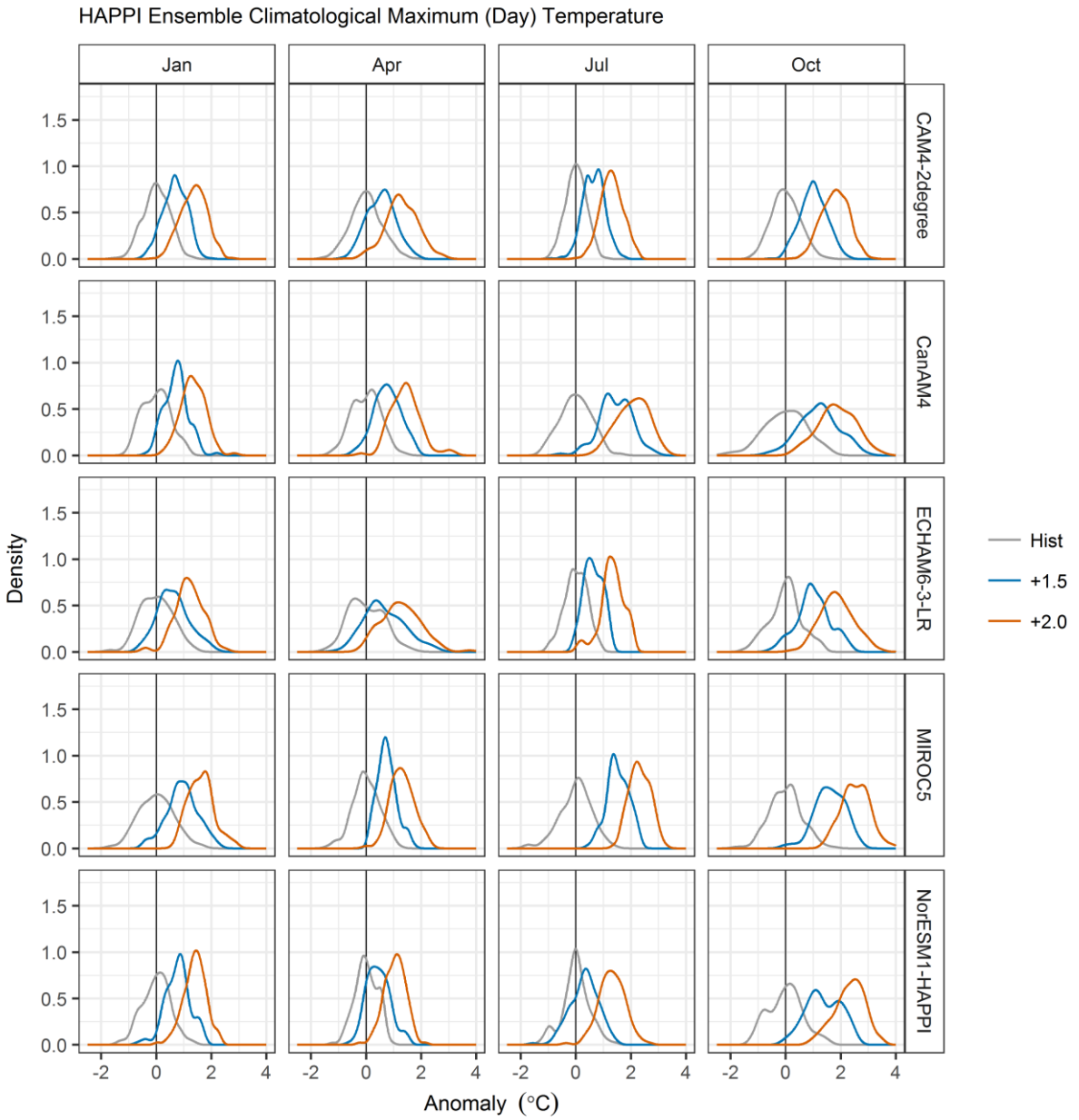


Figure 7.7 – As Figure 7.5, but for climatological monthly means of daily maximum (day-time) temperatures.

One possible explanation for the summer response is that summer forcing of minimum temperatures is strongly controlled by local influences and sensitive to representation of these influences in different models. Yet, interestingly, maximum temperatures do not show such pronounced differences between the distributions in different months (Figure 7.7). There is less spread in winter and more in summer compared with the simulated minimum temperature distributions. The reasons for this diurnal asymmetry in the spread of climatological temperatures due to internal variability are not immediately apparent, but it may be worth investigating further in future studies as a possible guide to seasonally differentiated predictability of possible climate responses.

CMIP5 Comparison

Figure 7.8 compares changes in the monthly climatology in a warmer world in the HAPPI model ensemble means with CMIP5. Only the 2°C warming scenario is shown, as the primary inferences are similar if the figure is reproduced with the 1.5°C scenario. Change factors are plotted by model, with separate lines for RCP4.5 and RCP8.5 realisations of the CMIP5 models. For precipitation, this indicates that both HAPPI and CMIP5 show relatively little change in the warmer scenario on average. There is notable variability in CMIP5 change factors between models/runs, which are largely confined to a band of +/- 25% of the historical baseline. However, there does not appear to be a coherent seasonal pattern within those bands, so much as a noisy response. For both CMIP5 and HAPPI, visually there could be a suggestion of increasing variability in the ensembles in late summer and early autumn, but it is unclear whether this is significant. There are some relatively extreme responses that move well outside the range occupied by most model runs. Inspecting the individual models (Appendix D, Section D.3) suggests that the more extreme precipitation increases are mainly associated with GFDL-ESM2G and MIROC5, while the largest reductions are associated with the IPSL-CM5A models.

For temperature, there is a clearer overall difference between the CMIP5 and HAPPI change factors, with the former being generally larger than the latter. This difference could pertain in part to the different historical baseline periods used out of necessity, as the two experiments cover different simulation periods. This could perhaps be partly addressed by reducing the size of the time window used in CMIP5 change factor calculations, but this might not be helpful given the limited number of runs with which to characterise internal variability relative to the HAPPI ensemble. There may also be some influence from the coupled models in CMIP5 representing a different type of experiment, with the resulting assumptions required to approximate a 2°C warmer world (Section 7.2.2).

It is also apparent from Figure 7.8b that the relative agreement of the HAPPI models decreases in late spring and early summer, driven by the MIROC5 model. There is some suggestion of higher change factors in late spring and autumn in CMIP5, but this is confounded by notable inter-model/run variation. There are some notably large and small change factors in late spring/early summer and winter, respectively, in the CMIP5 ensemble. Referring again to the change factors for individual models in (Appendix D, Section D.3), the

largest changes accompany the MIROC5 model, while the smallest are associated with FGOALS-s2.

This analysis therefore suggests that including both the HAPPI and CMIP5 ensembles in hydrological projections provides a reasonably large envelope of possible realisations of climate stabilisation in 1.5 and 2°C warmer worlds. Clearly there are multiple additional dimensions that could be considered here, including the effects of inter-annual and multi-decadal variability leading to relatively extreme wet/dry and cold/warm periods of relevance to water resources managers. Such variability will of course be highly significant even in notionally stabilised climate scenarios. Further investigation of such topics would usefully build upon the analysis here, which centres primarily on shifts in climatological mean states.

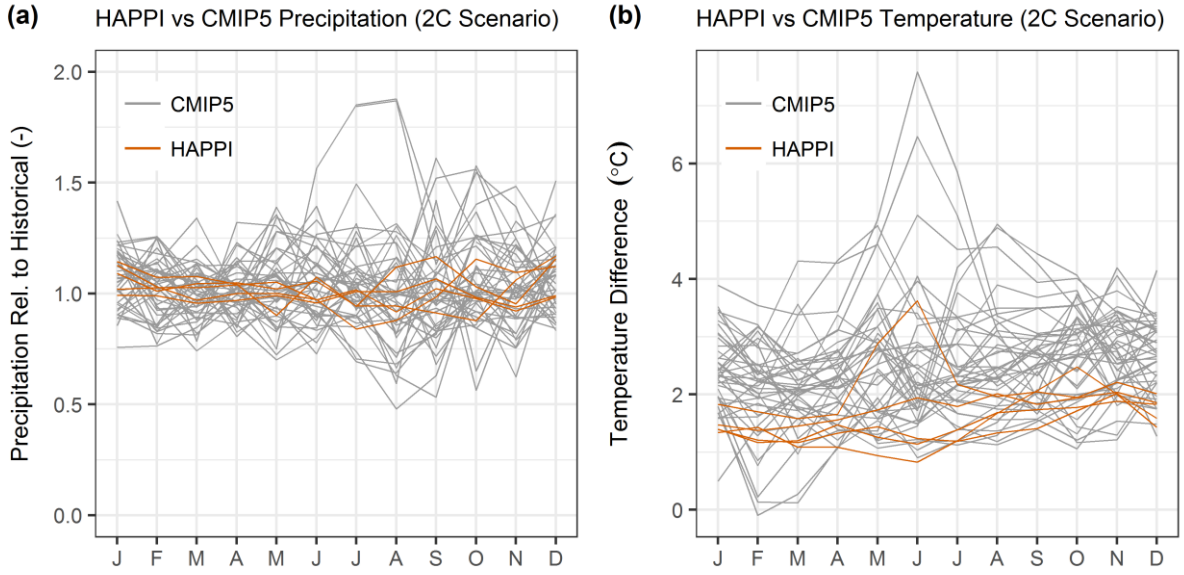


Figure 7.8 – Comparison of CMIP5 and HAPPI change factors for a 2°C warmer world. All models and realisations available are plotted as individual lines.

7.3.3 Hydrological Projections

The analysis of hydrological projections begins focusing on HAPPI climate scenarios using TOPKAPI-ETH. This is followed by a comparison of HAPPI and CMIP5 scenarios, again using TOPKAPI-ETH, which leads into an analysis of shifts in the annual water balance and flow components. Finally, a brief investigation of the influence of hydrological model choice is undertaken.

HAPPI Initial Conditions Ensemble Means

Figure 7.9 shows annual hydrographs for the Astore (a) and Hunza (b) sub-basins simulated by TOPKAPI-ETH for the baseline, 1.5 and 2°C warming scenarios. The climate scenarios applied in this case are monthly climatological change factors, with one set of runs using temperature perturbations only (experiment 1 - see Table 7.2) and a second set using both precipitation and temperature perturbations (experiment 2 - see Table 7.2). Beginning with the snow-dominated Astore sub-basin, Figure 7.9a shows a general tendency for the rising limb of the annual hydrograph to rise faster in the late spring and early summer in both 1.5 and 2°C warming scenarios compared with the baseline. The 2°C scenario leads the 1.5°C scenario in this shift in river flow towards the earlier part of the melt season, although the difference between the two scenarios is reasonably small.

There appears to be reasonable similarity in response between the GCMs, with the exception of MIROC5, which shows the most pronounced shift in the hydrograph. This fits with it exhibiting the highest temperature change factors in late spring and early summer (Section 7.3.2). MIROC5 is also the only GCM showing a notable increase in peak flows accompanying the change in timing. Moreover, the differences between experiment 1 (temperature perturbation only) and experiment 2 (temperature and precipitation perturbation) are minor. This reflects the limited change signals in precipitation in the means of the initial conditions ensembles for each model, such that temperature signals dominate the modelled response.

For the heavily glaciated Hunza sub-basin, Figure 7.9b shows some similarities for the patterns associated with the Astore in Figure 7.9a. These include the general tendency for river flows to rise more rapidly in the early part of the melt season compared with the baseline. However, with the exception of the simulations driven by MIROC5, this does not lead to the annual peak flow shifting forward notably. There are varying degrees of decline in river flows in the later part of the melt season in Figure 7.9b, when glacier melt is expected to become a particularly key component of river flows. Again, MIROC5 stands out as the model exhibiting the largest shifts. The large reduction in August runoff when using the MIROC5 future climate scenarios implies a substantial loss of glacier mass, as energy inputs are not likely to constrain glacier melt in this month. For all GCMs, the 2°C leads to both earlier runoff and lower late-season flows than the 1.5°C scenario, although for some models the differences relative to the baseline appear fairly small – particularly NorESM1.

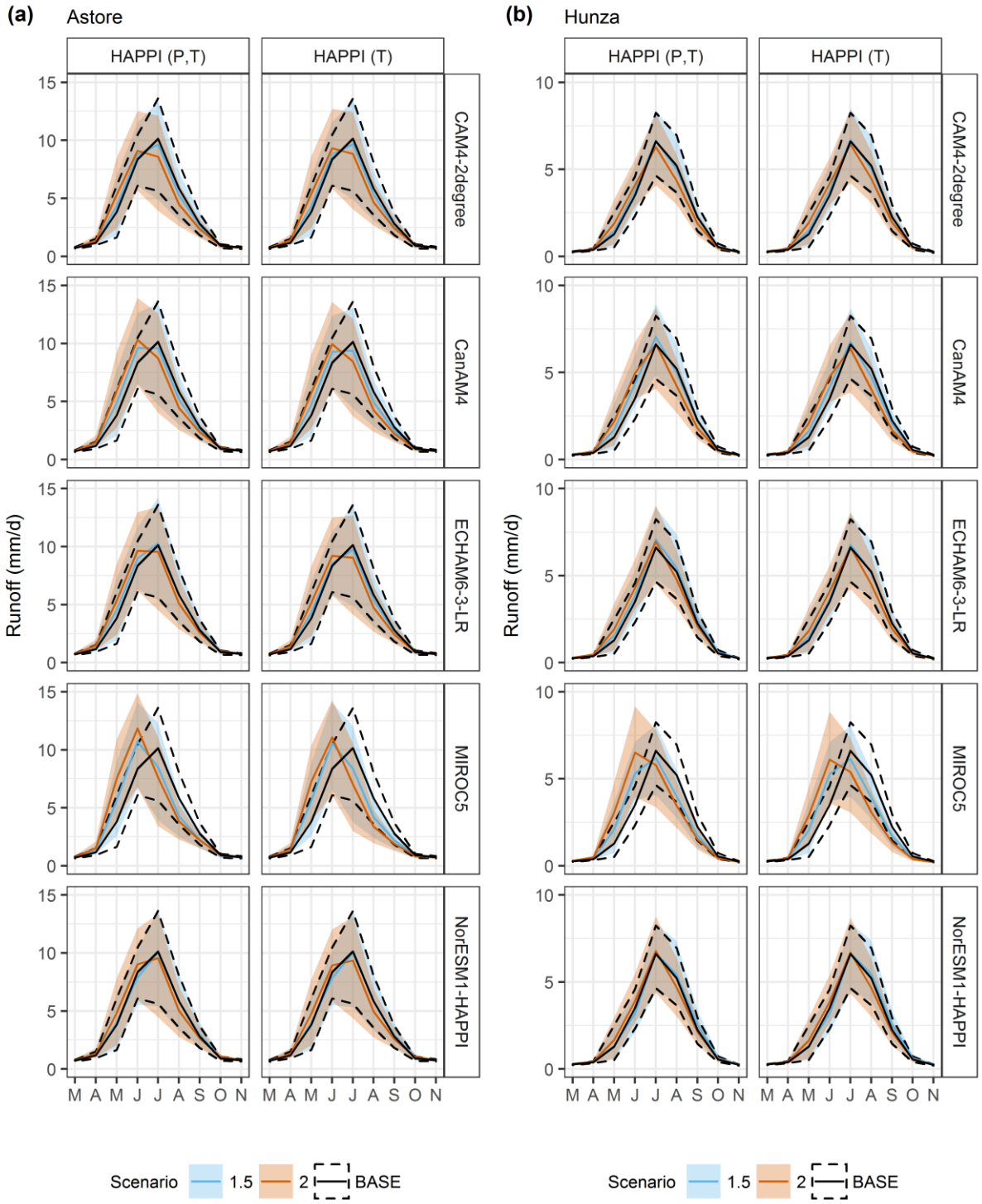


Figure 7.9 – Simulated annual hydrographs for historical baseline (BASE) and warming scenarios based on HAPPI models (change factors calculated from initial conditions ensemble means). Responses from perturbing precipitation and temperature (P,T) and temperature only (T) are shown in the left and right columns for each sub-basin, respectively. Lines and shading show means and 10th-90th percentile ranges for monthly flows, respectively.

HAPPI Change Factor Sampling

To investigate the influence of internal climate variability on possible hydrological changes, we now consider the groups of runs conducted based on sampling from the HAPPI initial conditions ensembles for each model. Corresponding with experiment 3 in Table 7.2, this sampling is based on using random pairs of runs from the baseline and warming scenarios to calculate change factors. Perturbations are applied for both precipitation and temperature. Further details of the method are given in Section 7.2.2.

Figure 7.10 shows the range of responses in mean flows each month from adopting this sampling approach to change factor calculation. The ranges depicted for the two warming scenarios correspond to the interval between the 10^{th} and 90^{th} percentile of the distribution of monthly mean flows. For context, inter-annual variability is shown for the historical baseline scenario, again using a 10^{th} to 90^{th} percentile range. Figure 7.10 suggests that the possible range of future mean flows tends to increase as the hydrograph peak is approached and decreases thereafter. In most cases this uncertainty remains lower than the inter-annual variability for historical flows, but for some models the uncertainty is particularly large. This includes scenarios driven by MIROC5, especially for the Hunza sub-basin, as well as scenarios forced with CanAM4 change factors. Uncertainty is relatively similar for 1.5 and 2°C scenarios, with some suggestion of slightly larger ranges for the latter for MIROC5 and CanAM4 at the height of the melt season.

In part, the ranges in Figure 7.10 reflect a methodological point regarding the significance of internal variability in obscuring changes in the mean annual hydrograph if only a small sample of climate model runs is available, or a small record period to define a baseline. To some extent Figure 7.10 also highlights the notable range of response associated with longer term shifts or variation (e.g. multi-decadal) in the mean runoff cycle under notionally "stable" climates in both the historical period and warmer worlds. It may well be worth further investigating approaches to sampling this variability. At any rate, the results do appear to highlight the importance of baseline and future climate definition from typically relatively short record periods even in the case of mean flows, before more extreme flow situations are taken into consideration.

Figure 7.11 is similar to Figure 7.10, except that it portrays uncertainty in the 10^{th} percentile of the (inter-annual) distribution of monthly flows, rather than uncertainty in the mean. The tendency for uncertainty to increase with progression towards the annual hydrograph peak in

summer is again apparent. Moreover, it is notable that the range of possible responses in the key summer months appear fairly large in relative terms, given the lower flows associated with the 10^{th} percentile of the distribution of flows in a month compared with the mean. This could have significant implications for water resources in dry years. However, there are of course limitations to this analysis, given that the projections apply monthly climatological change factors, without explicit consideration of inter-annual variability. Yet, as discussed in Section 7.2.2, this may still be an instructive sensitivity analysis at least, given the relatively small magnitude of change factor variation across the percentiles of the inter-annual distributions of both precipitation and temperature for most model/month combinations.

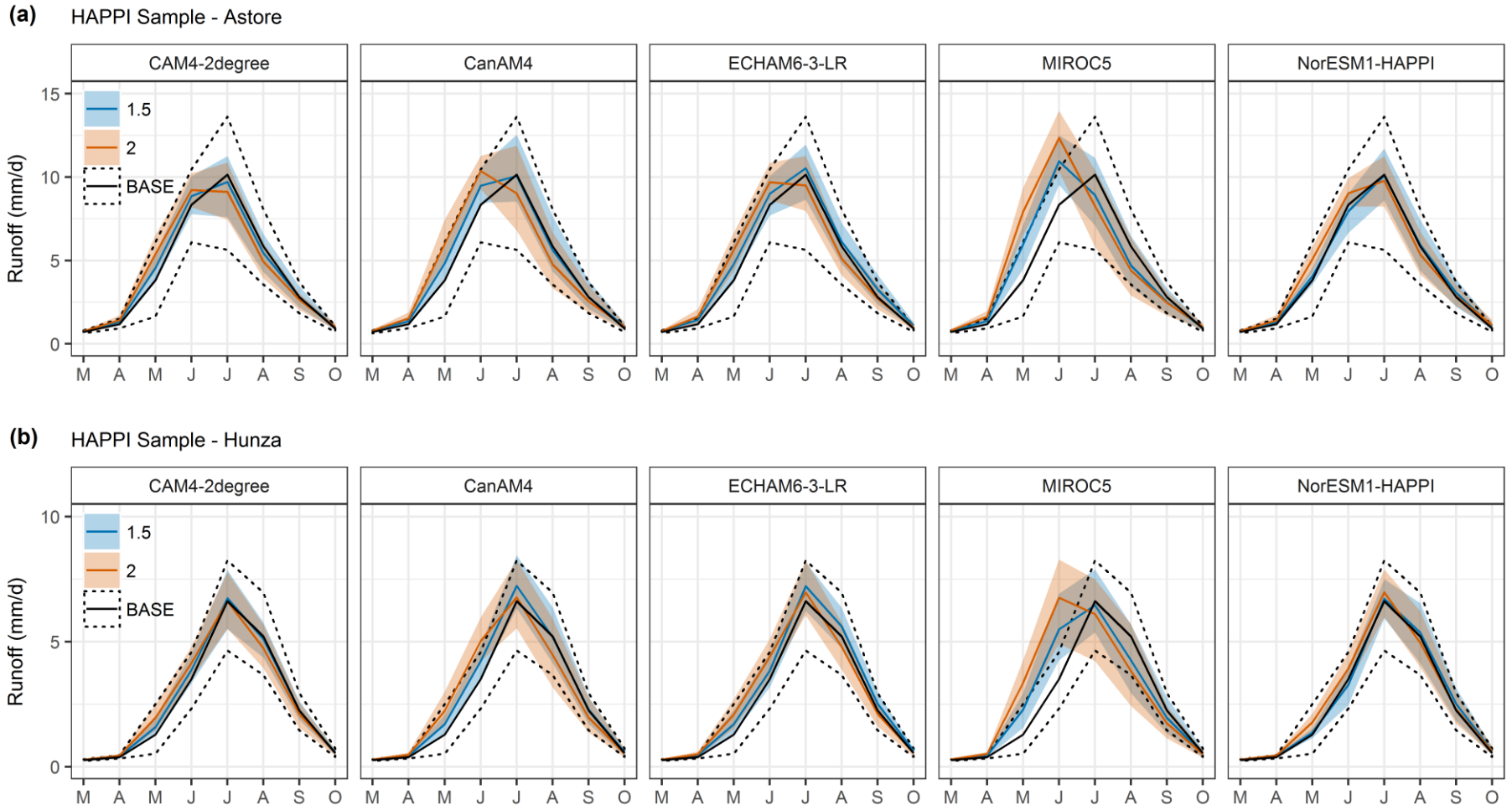


Figure 7.10 – Projections for monthly mean flows with TOPKAPI-ETH based on sampling from the HAPPI initial conditions ensembles for each model. The 10-90th percentile range and mean are shown for the HAPPI samples, with the baseline run showing inter-annual variability and mean runoff for context.

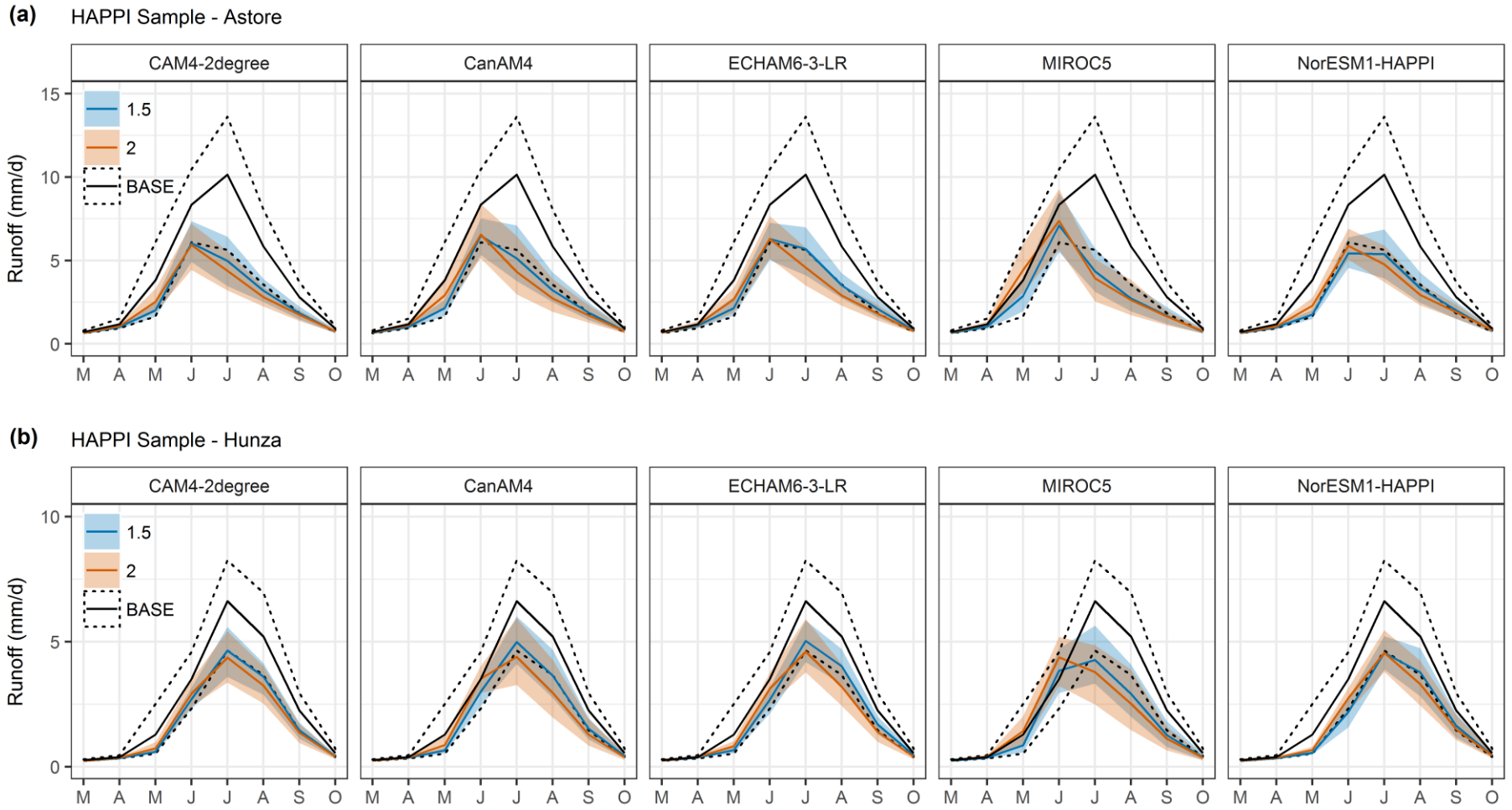


Figure 7.11 – As Figure 7.10 but for the 10th percentile of the inter-annual distribution of monthly flows for the future scenarios.

HAPPI Annual Flow Changes

Figure 7.12 shows the distributions of mean annual flow according to simulations based on sampling from the HAPPI initial conditions ensembles, as described above. This indicates that TOPKAPI-ETH driven by perturbations from HAPPI leads to a moderate increase in mean annual flow on average. The magnitude of this increase varies a little between models, but averages around 5-10% in the model/scenario combinations. There does not appear to be a consistent difference between 1.5 and 2°C scenarios in the mean annual flows in the model ensemble. Rather, most models show relatively minor differences, with the exception of MIROC5, which exhibits a more marked increase in flows under the 2°C warming scenario. This would seem to be driven by its larger increases in winter precipitation relative to the other models (Section 7.3.2). These patterns tend to hold for both the Astore and Hunza sub-basins, although it is possible that the contribution from shifts in climatological precipitation is primarily noise rather than a robust signal.

In most cases, the spread in the box plots in Figure 7.12 is fairly similar between models and 1.5/2°C scenarios, albeit with some reasonably small offsets between models. For most models the box (interquartile range) tends to sit wholly or largely above the baseline flow. This is associated with mean annual flows that range from around -10% up to +15% of baseline flows. In all models/scenarios, the boxplot whiskers (interquartile range multiplied by 1.5) straddle the baseline flow. In general the whiskers extend to around +/-25% of the baseline mean annual flows. This reaffirms the possibility that natural variability can strongly affect hydrological projections in climate stabilisation scenarios if there is no (or limited) means of sampling internal climate variability.

The HAPPI ensemble therefore suggests that there could be a shift towards higher flows earlier in the melt season under warming scenarios, as well as reduced flows later in the melt season. Sampling from the initial conditions ensembles suggests that changes in annual river flows may be fairly minor when averaged over sufficiently long periods, but it is also likely that variability due to inherent climate uncertainty may be quite significant even for mean flows over periods of a few decades. Hydrological projections based on HAPPI thus suggest that the primary issue for water resources managers in the Indus basin may be how to manage such variability in conjunction with a shift in flow timing, potentially including a concentration of flow into a potentially slightly shorter and earlier period.

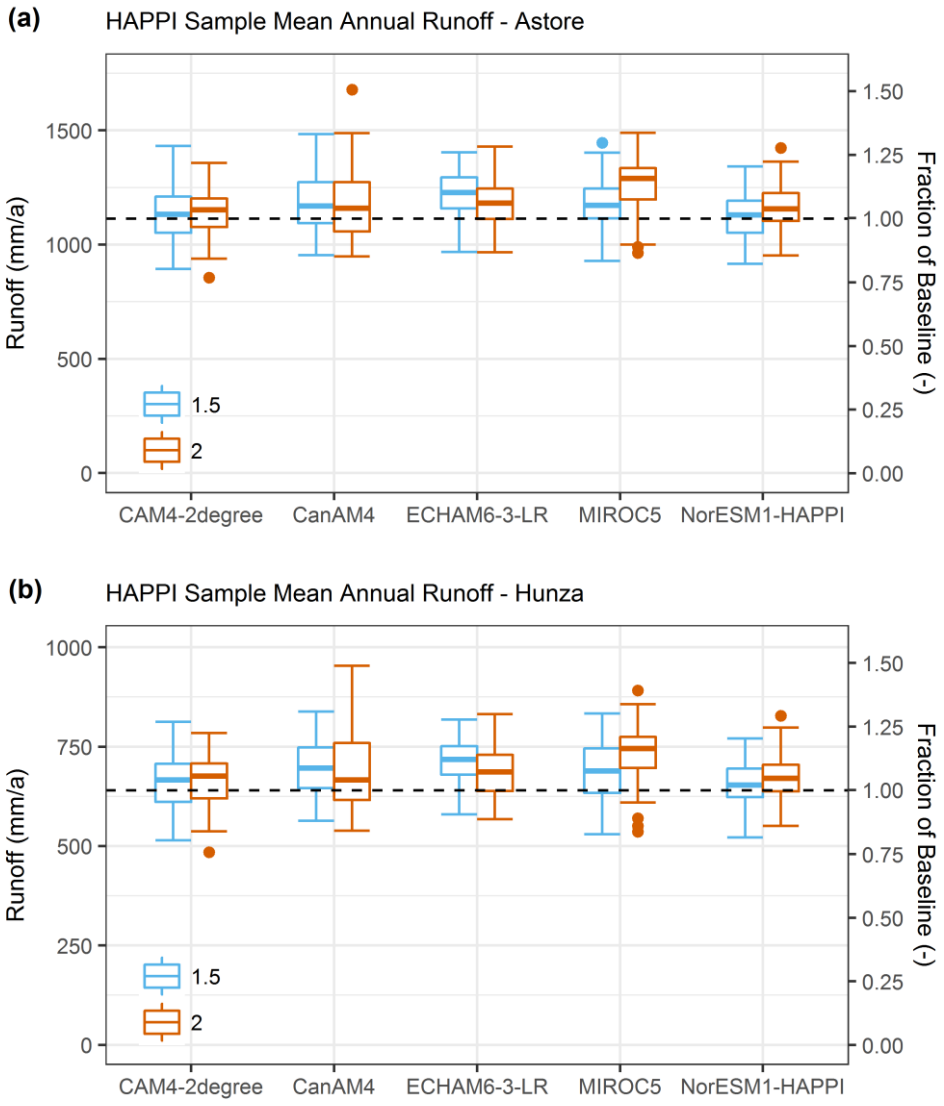


Figure 7.12 – Distribution of mean annual flows based on sampling from the HAPPI initial conditions ensembles to derive a range of change factors.

Comparison of HAPPI and CMIP5

Figure 7.13 compares HAPPI and CMIP5 in terms of hydrological projections. Specifically, the HAPPI climate scenarios are from experiments 1 and 2 (using means from the initial conditions ensembles to derive change factors), with the CMIP5 scenarios derived from outputs from both RCP4.5 and RCP8.5 (Section 7.2.2). For the future scenarios, Figure 7.13 shows the mean annual hydrograph for each simulation, but for the baseline scenario inter-annual variability is also displayed to contextualise the results. Beginning with the 1.5°C scenario and temperature perturbations only, both the CMIP5 and HAPPI ensembles show fairly good inter-model agreement with respect to changes in the annual cycle of monthly mean river flows. In effect the responses almost appear to fall into two groups, with most

models staying fairly close to the historical baseline but a small number of models showing a steeper rising limb and an earlier peak. This seems to apply to both the Astore and Hunza sub-basins, although the former possibly shows more CMIP5-driven ensemble members and a slightly higher peak flow in its group of runs exhibiting larger shifts from the baseline. In general the runs driven by HAPPI models also show smaller deviations from the baseline, especially for the Hunza, with the exception of the runs driven by MIROC5. This fits with the higher temperature change factors in CMIP5 evident in Figure 7.8.

For the 2°C warming scenario, again with temperature perturbations only, there is a further shift of the hydrograph into the earlier part of the melt season (Figure 7.13c and d). This results in more notable deviation from the baseline relative to the 1.5°C scenario. There is also a suggestion that the spread in the CMIP5 ensemble begins to increase further. For both the Astore and Hunza, there appears to be a group of CMIP5-driven runs (represented by the highest density of grey lines) that show relatively similar responses, which again lead the HAPPI-driven runs in the hydrograph response. However, it is also apparent that the groups of CMIP5 runs showing the most substantial deviations from the baseline rise more rapidly in the early part of the melt season relative to the 1.5°C scenario. This leads to a more rapid drop-off in flows in the middle and later parts of the melt season, which suggests some increase in the total ensemble spread, especially for the Hunza sub-basin. The MIROC5-driven HAPPI runs again deviate from the rest of the HAPPI ensemble, which show notable consistency. This may indicate that the magnitude of HAPPI inter-model differences in temperature change factors portrayed in Figure 7.3 and Figure 7.4 are generally of relatively minor importance for hydrological response, with the main exception of the very pronounced May/June temperature change factors in MIROC5.

Figure 7.13e and f demonstrate that the complexity and divergence in hydrological response increases yet further if precipitation perturbations are accounted for, in addition to temperature. This is particularly the case for CMIP5. While the fundamental shifts in hydrograph timing are largely similar to the temperature-only perturbation case, the overall range of responses increases and the relatively coherent groupings in CMIP5-driven runs break down to a degree. There is also some more modest increase in spread in the HAPPI-driven runs, more so for the Astore sub-basin. Yet, as only two realisations of future climate were used for the CMIP5 models (associated with RCP4.5 and RCP.5), the results again highlight the possibility of substantial hydrological implications arising from internal variability governing precipitation change factors in CMIP5 in the UIB. Where the large

initial conditions ensembles are available for HAPPI, the earlier results suggest that very long-term precipitation shifts in the warmer worlds may not be wholly detectable in the UIB. However, the variability shown in the CMIP5 responses, and indeed when sampling from the HAPPI initial conditions ensembles, may provide the more useful guide to the range of potential realisations of future climate to be accounted for in water resources planning. Clearly some models may be legitimately excluded based on process-based considerations, such as those in Chapter 6, but we leave it for additional detailed analysis to consider the many dimensions of this problem.

Figure 7.14 provides a similar analysis to Figure 7.13, except this time for the 10th percentile of monthly flows, rather than the mean. This provides an initial exploration of potential changes in lower flow conditions in warmer worlds, with the associated implications for water resources availability. In general, Figure 7.14 suggests that the fundamental shifts in flow timing and the structure of variability in simulated response are similar for the 10th percentile of monthly flows as described above for mean flows. Particularly in the 2°C warmer world, the hydrograph rises more steeply in the early part of the melt season, before declining more rapidly in the middle and later part of the season. In relative terms, these changes can be very substantial. Moreover, for the Hunza there appears to be a particularly large spread in simulated CMIP5-driven responses, especially when both precipitation and temperature changes are taken into account. This likely pertains to the particularly high proportion of total runoff supplied usually by glacier melt in comparatively dry (lower flow) years. The results imply that the glacier mass balance trajectories in a 2°C world under CMIP5 are highly variable, such that the "buffering" function of glaciers in drier years is highly uncertain – but in some cases very substantially reduced.

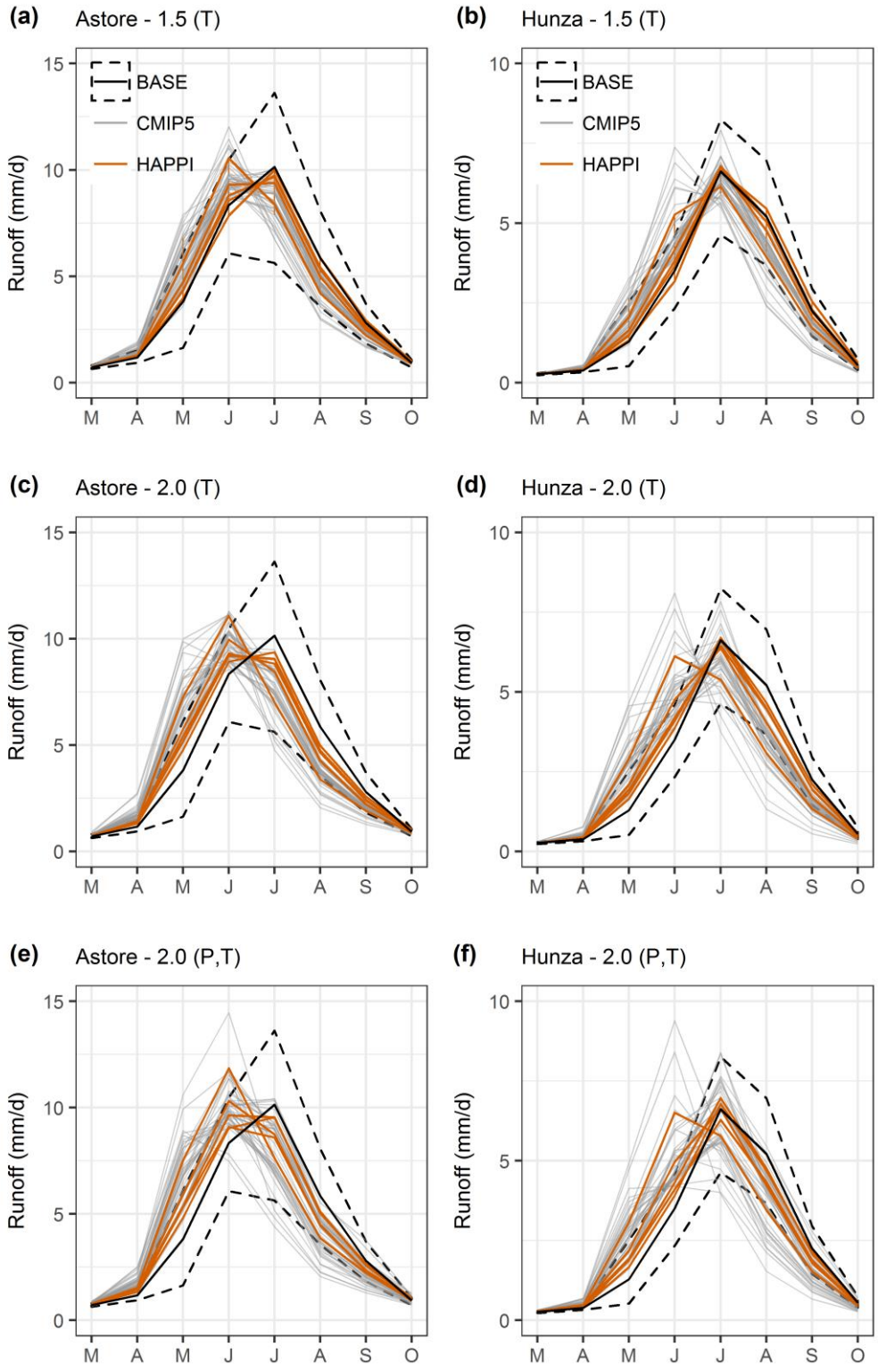


Figure 7.13 – TOPKAPI-ETH monthly mean river flows using CMIP5 and HAPPI change factors (P – precipitation, T – temperature). Initial condition ensemble means are used for HAPPI change factor calculations, while CMIP5 change factors are calculated separately using RCP4.5 and RCP8.5 runs. For the baseline scenario, the mean and 10-90th percentile range of monthly flows are shown.

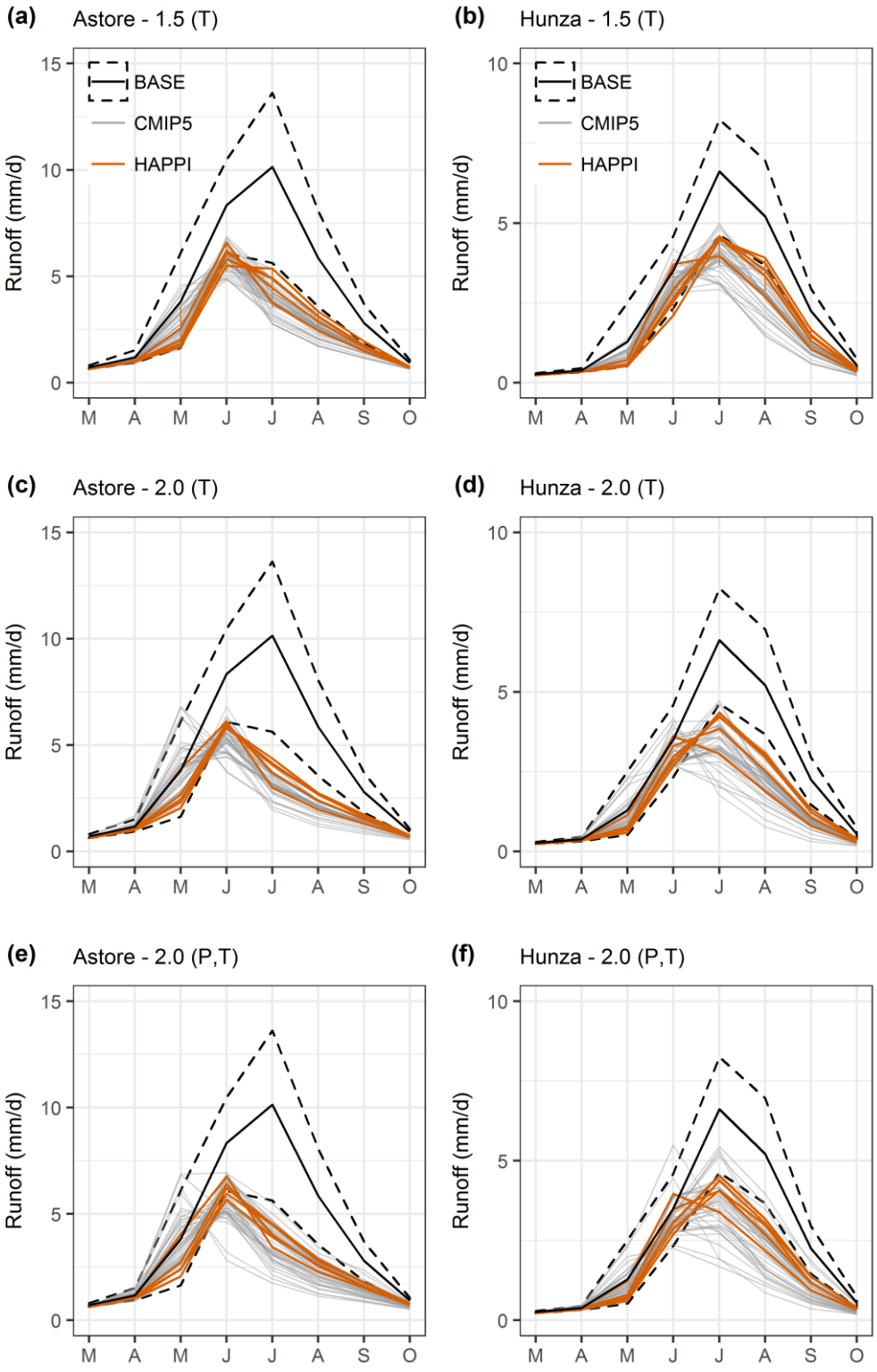


Figure 7.14 – As Figure 7.13, but for the 10th percentile of monthly flows, rather than the mean.

Comparison with Higher Global Temperature Scenarios

As HAPPI scenarios are not available for global warming scenarios greater than 2°C, we use CMIP5 outputs to provide an initial indication of the implications of higher temperature changes. As an initial exploration of this we focus on overall water availability. However, a preliminary comparison of the annual cycle of river flows based on CMIP5 scenarios for 2 and 3°C warmer worlds for the Astore and Hunza sub-basins is provided in Appendix D (Section D.4). The primary finding from this is that the rising limb and peak of the annual hydrograph are brought forward yet earlier in the year under a 3°C scenario in both flow regimes, while flows later in the melt season decline notably in most cases. Again, however, there is notable variation in the extent of change relative to the 2°C scenario depending on which CMIP5 model provides change factors. In most cases, there is in fact quite good agreement between the simulated responses when deriving change factors from different RCPs, especially when just temperature is perturbed. Including precipitation complicates the response, particularly after the annual hydrograph peak is reached and catchment response potentially becomes mass- rather than energy-limited.

To explore the implications of 1.5, 2 and 3°C warming scenarios for overall water availability, Figure 7.15 and Figure 7.16 consider annual flows in the TOPKAPI-ETH hydrological projections driven by CMIP5 for the Astore and Hunza sub-basins. The scenarios use both precipitation and temperature perturbations. The first point to note is that, for the CMIP5 ensemble as a whole, there is little overall change in mean annual flows (of the order of a few percent relative to baseline flows) in each of the warming scenarios shown (panel a in Figure 7.15 and Figure 7.16). However, there is substantial inter-model variation, which leads to differences of up to -20% to +40% in a 3°C scenario relative to the baseline. This pattern of inter-model variation is similar for both sub-basins. Moreover, in the case of the drier conditions implied by the 10^{th} percentile of the distribution of annual flows (panel b in Figure 7.15 and Figure 7.16), a shift in water availability starts to become more noticeable for the ensemble as a whole in a 3°C scenario. This reflects the erosion of glacial storage referred to above, which is particularly relevant to the Hunza sub-basin and in drier periods. For some models in the 3°C scenario, reductions in annual water availability of over 25% are apparent in drier conditions.

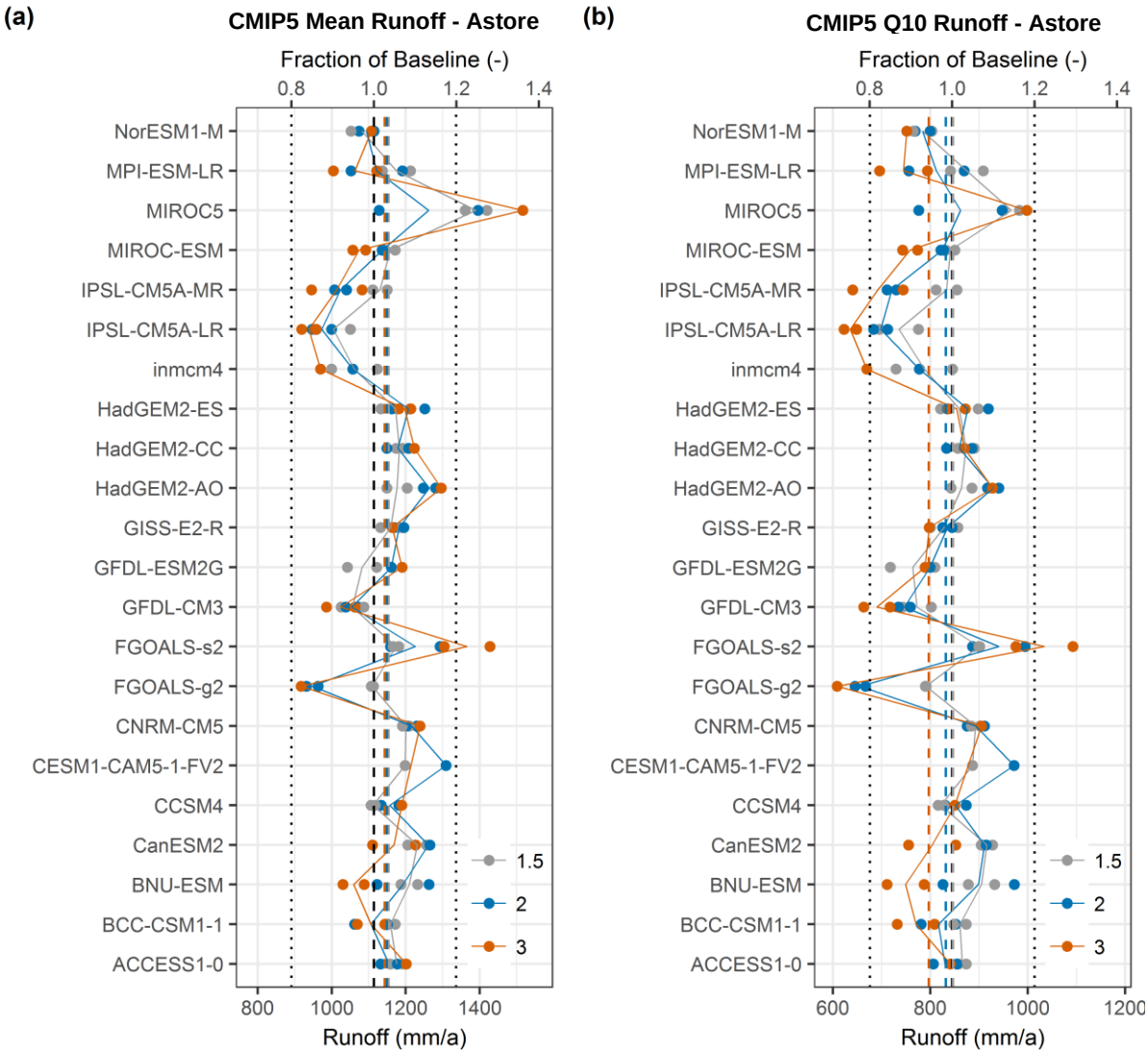


Figure 7.15 – Annual flows in TOPKAPI-ETH driven by CMIP5-derived scenarios for the Astore sub-basin. Panel (a) shows mean annual flows, while panel (b) shows the 10th percentile of the distribution of annual flows (i.e. from inter-annual variability). Dashed lines show the ensemble mean change for each scenario, while dotted lines just indicated +/- 20% relative changes to help with interpretation.

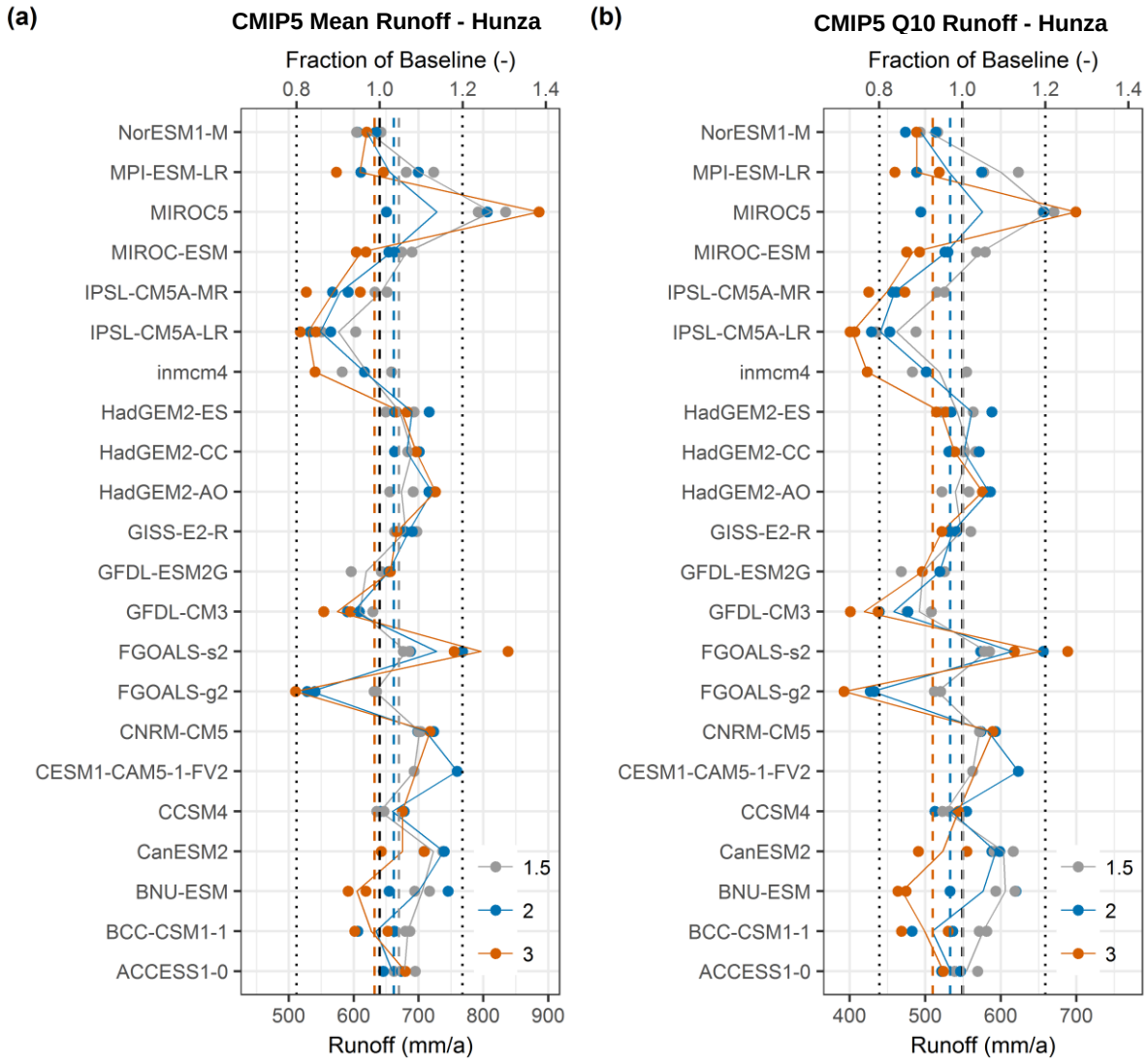


Figure 7.16 – As Figure 7.15 but for the Hunza sub-basin.

To understand more fully the annual water balance implications of these various realisations of UIB climate in warmer worlds, Figure 7.17 and Figure 7.18 show the associated changes in mean annual precipitation and evapotranspiration. The key point here is the very strong similarity between the precipitation profiles in Figure 7.17 and the runoff profiles in Figure 7.15 and Figure 7.16. This suggests that the inter-model variation in precipitation change factors dominates the variation in annual flow response. Evapotranspiration changes tend to be relatively coherent between models and proportionally larger than precipitation changes on average, but the absolute changes simulated by TOPKAPI-ETH seem to play a notably secondary role in shaping gross water availability under average conditions. This suggests that, for the mean annual case, temperature projections are primarily implicated in the nature and magnitude of seasonal shifts in flow timing. These shifts likely embody progressive transitions in flow regime, such that a strong glacial component gives way to increasing

seasonal snow melt, which could ultimately be followed by increasing prominence of rainfall in comparatively lower elevation sub-basins (Lutz et al. 2016a).

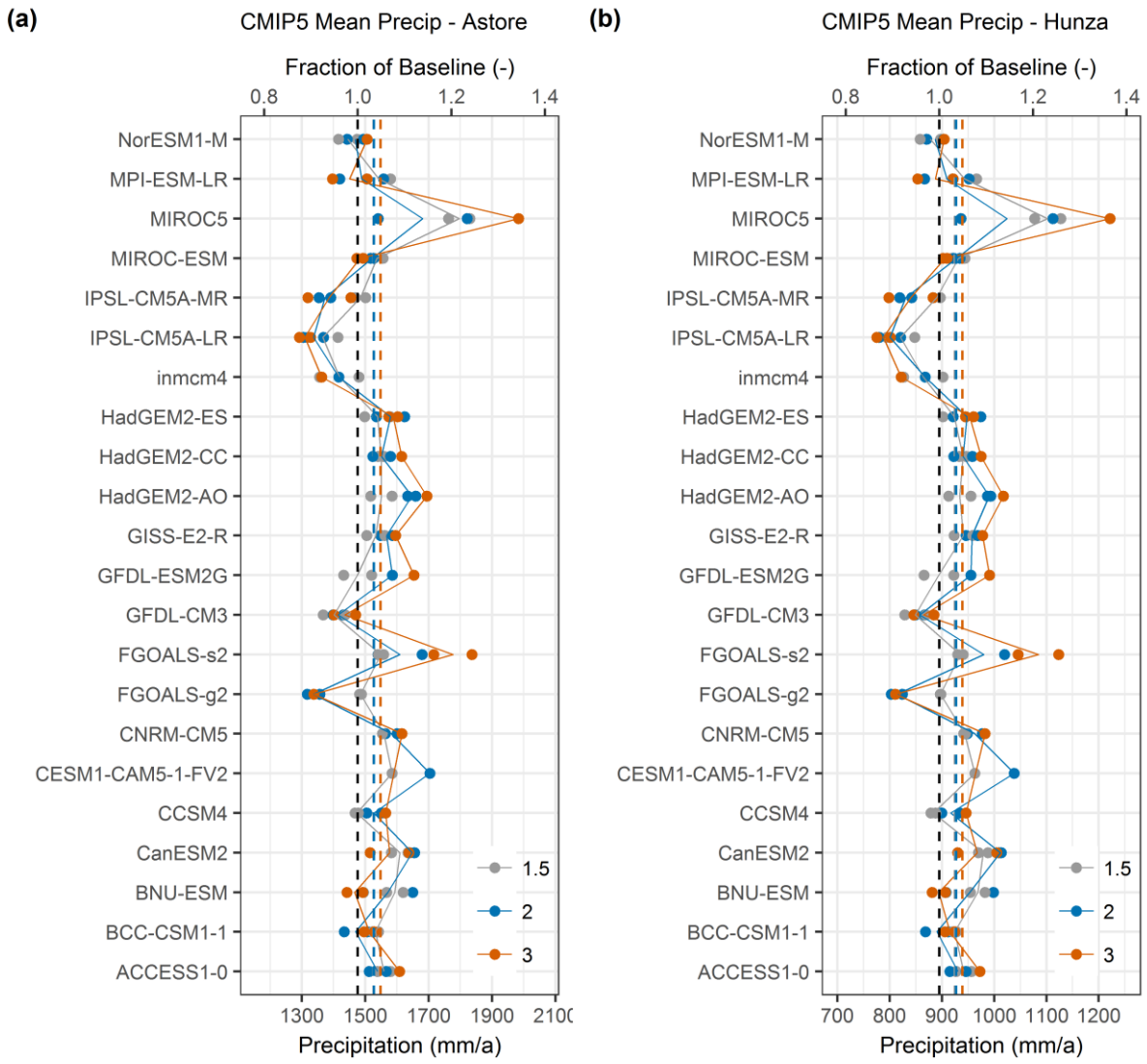


Figure 7.17 – Similar to Figure 7.15 but for mean annual precipitation, and with both the Astore and Hunza sub-basins shown together.

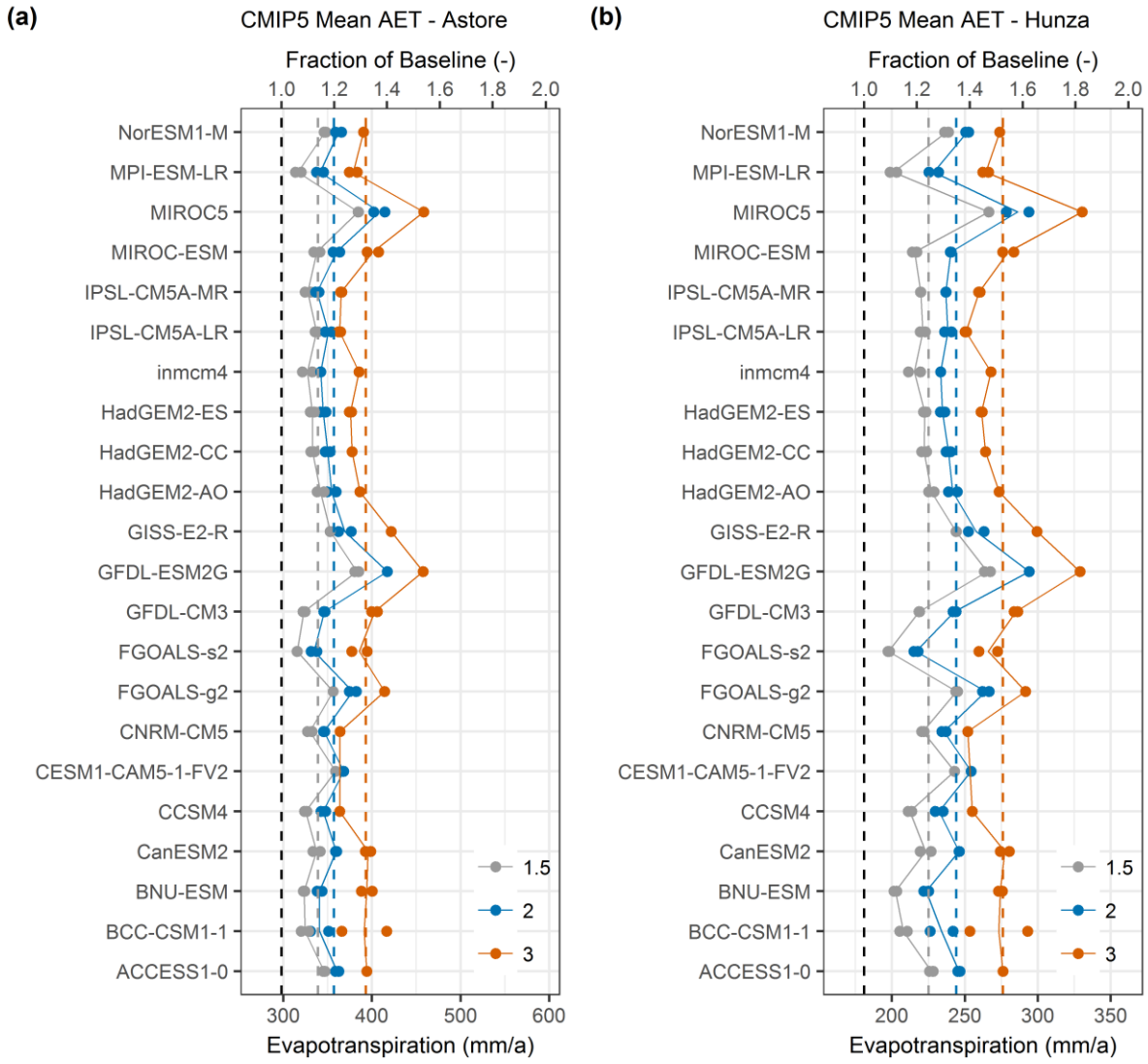


Figure 7.18 – As Figure 7.17 but for mean annual actual evapotranspiration.

However, as discussed above, the apparent significance of temperature begins to change in dry/low flow periods. Figure 7.19 reflects this point, showing that there are significant decreases in glacier contribution to flows in all of the warming scenarios, but very large inter-model variation in the 1.5 and 2°C scenarios relative to the 3°C scenario. This variation is closely related to the magnitude of temperature change in different CMIP5 models, as evidenced by the anti-correlation with Figure 7.18. Interestingly, Figure 7.19 shows that there is much more uncertainty (inter-model variation) in the contribution of glacier melt to mean annual flow in the 1.5 and 2°C scenarios compared with the 3°C scenario. This suggests that, under low or moderate warming scenarios, the extent of simulated glacier wastage varies strongly with the different change factors in the climate model ensemble. By the time higher scenarios are reached, more ubiquitous glacier mass loss is consistently simulated. There is a strong possibility that variations in the representation of snow, albedo and temperature

feedbacks in the climate models are implicated in the larger inter-model differences under 1.5 and 2°C scenarios.

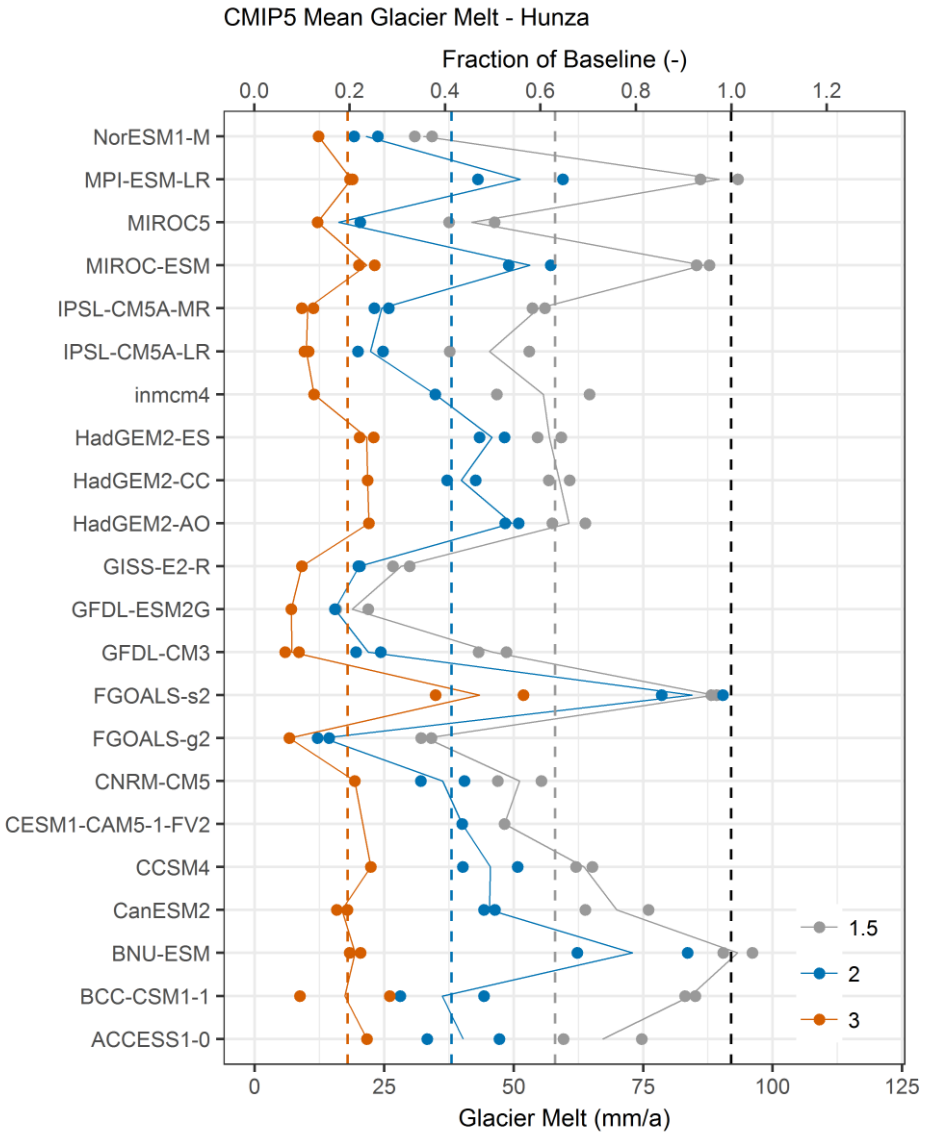


Figure 7.19 – As Figure 7.17 but just for the Hunza and for mean annual glacier melt.

7.4 Conclusions

This study provides the first analysis of climate and hydrological impacts in the UIB at 1.5°C, 2°C and higher global warming stabilisation targets using large ensembles of climate model simulations. In terms of climate impacts, the HAPPI models suggest that there is a local or regional amplification of the global warming signal in the UIB in 1.5 and 2°C warmer worlds. Indeed, the difference between the 1.5 and 2°C scenarios tends to exceed the half degree difference in global temperatures – on average by 0.7-0.8°C, but by up to 1.4°C. Interestingly, the distributions associated with the initial conditions ensembles seem to show different levels

of uncertainty in projected temperature changes at different times of year. This is particularly the case for minimum (night-time) temperatures, where internal variability appears to be very important for climatological temperatures in winter/spring while having very limited influence in summer. On average, temperature increases derived from CMIP5 tend to exceed those in HAPPI, acknowledging notable within-ensemble variation in the former. This may be at least partly due to the different historical reference periods necessitated by the differences in the experiment designs. While Chapter 6 raises questions about whether UIB warming rates in climate models may be too rapid, owing to biases in snow-albedo feedbacks and other issues, the range of projected temperature rises in HAPPI and CMIP5 models do at least provide multiple possibilities to consider for adaptation. From a planning perspective, enhancing resilience to various warming rates is likely to represent a precautionary approach, given that notable uncertainty may persist for some time.

For precipitation, the distributions associated with the initial conditions ensembles for HAPPI suggest that there may not be much clear difference in mean precipitation between the historical and future scenarios. However, different realisations of the HAPPI runs for a given model show notable variation in climatological precipitation. This suggests that internal variability may play a key role in water availability over inter-annual to multi-decadal time scales in 1.5 or 2°C warmer worlds, perhaps more so than forced response. The large initial conditions ensembles in HAPPI may also help to contextualise the notable variation in projected precipitation changes between CMIP5 models. Taken as a whole, the CMIP5 ensemble shows little coherent change in precipitation in the 1.5 and 2°C warming scenarios. When viewed together with HAPPI, this may be in large part due to natural variability, although a role for model structure cannot be ruled out at this point.

In the notionally stabilised climates associated with 1.5, 2 and 3°C warmer worlds, successively larger shifts in flow timing are simulated by TOPKAPI-ETH. Flows tend to increase in the early part of the melt season, with earlier onset in melting evident in the 3°C scenario particularly. Conversely, simulated flows decrease in the later part of the melt season. However, when sampling from the HAPPI initial conditions ensembles (to get a range of change factors) and considering the full CMIP5 ensemble, the uncertainty in hydrograph shifts can become relatively large compared with the magnitude of historical inter-annual variability. This variability in hydrological response increases particularly when precipitation is taken into consideration. With temperature perturbations alone, there are stronger groupings and lower spread in the simulated responses at sub-basin scales.

The results do not suggest a great deal of change in mean annual water availability from CMIP5 or HAPPI ensemble means. On average, the rising flows in the earlier part of the melt season are thus largely offset by decreasing flows in the later part. However, for drier years this begins to change under the 2 and 3°C warming scenarios, particularly for the Hunza sub-basin where the buffering effect of glacier melt gets eroded by ice mass loss. More important perhaps is the variability in water availability in mean and drier conditions associated with natural climate variability. Exemplified by the CMIP5 ensemble, precipitation changes are the primary control on changes in annual water availability in the climate stabilisation scenarios. As noted above, this suggests that natural variability in precipitation over a range of time scales is likely to be critical to the apparent hydrological response over the coming decades. The increases in evapotranspiration in warmer worlds have less of an effect in the simulations, although there may be some degree of influence from the TOPKAPI-ETH evapotranspiration and soil moisture balance algorithms here. This could be investigated further. Therefore, the fact that quite a large range of precipitation possibilities are projected by both CMIP5 and the HAPPI initial conditions ensembles suggest water management needs to account for variability over inter-annual to multi-decadal time scales, in addition to the overall shift in flow timing and the possibility of greater reductions in flow during drier years/periods.

One of the key questions arising from this is whether the range of climate projections can be constrained. The results in Chapter 6 may ultimately provide a basis for this. Specifically, observational constraints based on snow albedo climatology and temperature feedbacks may be suitable for eliminating some implausible models. Further investigation may reveal additional process-based considerations with which to constrain ensembles of climate projections for regional impact studies. In parallel, it would also be useful to re-examine the projections to compare the differences between TOPKAPI-ETH and process-based models solving the surface energy balance. One key aspect of this will be whether the models diverge substantially at higher elevations, which would affect the long-term trajectories of the perennial cryosphere. However, investigating this will require improved downscaling methods to obtain physically consistent climate fields across multiple variables in complex terrain.

Chapter 8

Conclusions

Projections of cryospheric and hydrological changes in the Upper Indus Basin (UIB) are highly uncertain. Much of this uncertainty arises due to challenges in: characterising near-surface climate fields needed for model input; selecting appropriate model structures to balance process fidelity with data availability; and understanding the wide spread in climate model projections used in impact assessments. As such, this thesis has attempted to identify pathways for refined hydrological projections in the UIB. Five objectives were defined in Chapter 1 to help achieve this aim. The present chapter begins by summarising and discussing the main findings in relation to these objectives. This is followed by a discussion of possible further work.

8.1 Summary and Discussion of Results

Objective 1: To evaluate how well the High Asia Refined Analysis (HAR) represents near-surface climate in the UIB, with reference to its potential for improving climate inputs for hydrological modelling

With data paucity underpinning many of the challenges in modelling the hydrology of the UIB, it is timely to consider how alternative products such as the HAR may be best employed. While the HAR has shown great potential in studies of climate and the cryosphere in other parts of the Himalaya and Tibetan Plateau (Maussion et al. 2014; Curio et al. 2015; Curio and Scherer 2016; Huintjes et al. 2015; Biskop et al. 2016; Tarasova et al. 2016; Mölg et al. 2014), it has not previously been evaluated in the UIB. Moreover, prior work has not placed its emphasis on biases in the climatology of multiple variables and how they relate to each other. Such analyses are necessary to determine the potential of model-based products like the HAR to support applications in the UIB. They also provide some insights into possible causes of deficiencies in the HAR, which may inform future coupled land-atmosphere modelling efforts in the region.

The evaluation in Chapter 3 shows that the HAR has both strengths and weaknesses in simulating UIB near-surface climate. One of the major strengths is its precipitation climatology. This is evidenced by its generally good consistency with multiple reference

datasets pertaining to different scales. These include in-situ point measurements and catchment-scale observed runoff, as well as vertical/spatial gradients inferred from observations and remote sensing. The HAR also exhibits reasonable consistency with in-situ observations for temperature, humidity and incoming radiation in many respects. However, some notable, seasonally varying biases are present. Perhaps most crucially, temperature exhibits a cold bias that worsens in spring but reduces in summer. This seasonal variation appears to be at least partially related to deficiencies in HAR cloud, snow and albedo representations affecting surface radiative balances and fluxes. Specifically, high albedo in spring may suppress surface warming, while low cloud cover in summer may induce overestimation of incoming shortwave radiation and thus surface heating.

On the issue of seasonally varying biases, the importance of snow and albedo representations in the Weather Research and Forecasting (WRF) model has recently been demonstrated in several other contexts (e.g. Tomasi et al. 2017; Meng et al. 2018; Saha et al. 2017). García-Díez et al. (2015) also noted that albedo issues may act mainly as a feedback amplifying poor representations of the snowpack or snow-atmosphere interactions, rather than a primary driver of temperature bias. Indeed, it may be that such a feedback underpins the large peak cold bias in spring in the HAR, possibly by reinforcing biases induced by other aspects of the WRF configuration. For example, the combination of snow process parameterisations with planetary boundary layer (PBL) schemes could be crucial (García-Díez et al. 2013). Thus, an important question raised by this work regards the extent to which WRF simulations for the UIB could be improved by revisions to snow initialisation, snow process representations and sub-grid variability parameterisations. Yet despite issues such as these, the results in Chapter 3 confirm that the methods underpinning the HAR show considerable potential for augmenting sparse observations and other data sources in the UIB.

Objective 2: To assess the feasibility of process-based snow modelling in the UIB, as well as the implications of alternative process representations

As the HAR shows a number of positive features with respect to local observations, remote sensing and other gridded data products, it opens up the possibility of testing the feasibility and skill of process-based cryospheric and hydrological models in the UIB. In the past this was largely inhibited by the dearth of physically consistent, high resolution climate fields needed for model input. In parallel with this, recent developments in modelling frameworks have created opportunities to better understand how different representations of processes

shape model responses and performance. Such frameworks have not been much applied in the UIB and neighbouring basins. As such, they offer a largely unexploited potential to understand modelling choices and improve conceptual understanding. This potential is heightened by the availability of multiple remote sensing products to augment local observations. The application of these frameworks is particularly key with respect to snow processes, which remain little studied despite their vast climatic and hydrological significance.

The results in Chapter 4 suggest that spatially distributed, process-based snow modelling for hydrological applications in the UIB is feasible with HAR-based climate inputs. The overall correspondence of snow model outputs with multiple reference datasets thus provides further verification of the HAR and its practical utility following minimal bias correction and spatial disaggregation. Critically, however, different snow model configurations lead to notable variation in model response and performance. For hydrological applications in the study area, the Factorial Snowpack Model (FSM) ensemble spread depends primarily on how processes related to albedo, liquid water in the snowpack and turbulent heat fluxes are represented. Propensities to accelerate or delay snowpack runoff are present in different parameterisations of these processes, which interact to exacerbate or compensate each other's tendencies. Based on observations and multiple remote sensing products, combining prognostic albedo with inclusion of snowpack hydrology generally leads to the best model performance. This is largely consistent with previous (e.g. Essery et al. 2013; Magnusson et al. 2015), although at the catchment-scale and with a hydrological focus we find a lower sensitivity to snow density parameterisations than in some other contexts.

The results also suggest that applying the atmospheric stability adjustment may suppress turbulent fluxes too much. The roles of input uncertainty and unresolved processes, such as local heat advection, may provide some confounding influence here, but this issue of excessive flux damping under stable conditions has been highlighted previously in various contexts (e.g. Jordan 1991; Tarboton et al. 1995; Andreadis et al. 2009; Conway and Cullen 2013; Collier et al. 2015). The issue is thus likely to be both substantive and persistent, given the paucity and difficulty of turbulent heat flux measurements in mountainous environments like the Himalaya (e.g. Stigter et al. 2018). Empirical corrections may be more feasible in the immediate term, such as the constant adjustment factor applied by Collier et al. (2015). This raises the question of how best to partition and use the observations and remote sensing available for calibration and/or evaluation.

The process options in FSM essentially behave and interact similarly between years, but no single configuration appears to perform the best in all years simulated. Rather, trade-offs in performance are apparent, with models that perform better in some years being worse in others and vice versa. This may reflect possible structural constraints on model performance, input uncertainties, or perhaps an undesirable and systematic dependence of model response on climate conditions. In many ways, these possibilities fit with issues raised in previous snow model inter-comparison studies. Using data from well-instrumented sites, these studies have generally found model performance to vary between locations and years, albeit with groups of configurations that tend to perform consistently well, poorly or variably (Etchevers et al. 2004; Rutter et al. 2009; Essery et al. 2009, 2013; Magnusson et al. 2015; Lafaysse et al. 2017). The UIB represents a case of much higher uncertainty in both input and evaluation data, but it is interesting that the catchment-scale inter-comparison undertaken with FSM raises similar issues to small-scale site-based analyses. The issues leading to spatial and temporal performance variation may thus be similar at larger scales.

In conjunction with potentially significant variations in climate sensitivity between FSM configurations, this suggests that an ensemble-based approach would be useful in applications such as climate change impact assessment. However, some configurations could be excluded. These would most likely be configurations using diagnostic albedo and omitting the liquid water drainage and retention parameterisation. Beyond hydrological applications, the results here raise questions about the snow process representations applied in different land surface models (LSMs) for regional or global climate modelling. As discussed above in relation to the HAR, deficiencies in snow process parameterisations may be strongly implicated in the shortcomings of coupled land-atmosphere modelling in the UIB and neighbouring basins. The results may therefore provide a partial guide to assessing which LSMs applied in different weather and climate models are more or less reasonable.

Objective 3: To compare process-based models of snow and glacier ablation with simpler approaches for simulating UIB hydrology

While this study shows that the scope for applying process-based models like FSM in the UIB is growing, some important questions remain regarding how they differ from more widely applied, simpler models. In part these questions pertain to differences in model performance. Perhaps more critical, however, is how contrasting model formulations might systematically affect hydrological projections. Simpler models often show very good performance relative to

observations, acknowledging the data limitations of the UIB, but fundamentally they need to show appropriate sensitivity to climate perturbations to provide reasonable hydrological projections. This is not necessarily straight-forward to deduce from historical performance alone, especially when evaluations are often unavoidably limited to short time periods relative to the memory and response times of cryosphere-dominated hydrological regimes. Therefore, there is a need to compare the functional form of model response in simple and more process-based models.

Focusing on ablation, the results in Chapter 5 suggest that there are indeed some notable differences in behaviour between model formulations of contrasting complexity. For snowpack modelling, parameters would need to be varied substantially for temperature index (TI) and enhanced temperature index (ETI) formulations to match the responses of a more process-based model like FSM. This variation is partly systematic, meaning that some consistent spatial and seasonal patterns of implied parameter variation are evident. However, there is also substantial inter-annual variability. While FSM is itself a simplification of reality, the degree of inter-annual variability raises the possibility that important feedbacks are missed in the simple formulations. This issue may become particularly pronounced at high elevations, where the relationship between temperature (through positive degree days) and annual melt totals begins to deviate from that observed over much of the elevation range. Representation of liquid water drainage, retention and melt/refreezing cycles in snow may be particularly critical in simulating the high elevation water balance appropriately, with associated implications for perennial snowpacks and glacial mass balance.

By additionally incorporating a term for net shortwave radiation, the ETI model shows more agreement with energy balance approaches like FSM than the TI model. However, using FSM outputs to deconstruct the ETI formulation (Pellicciotti et al. 2005) suggests that its weaker performance under cloudy conditions may be a structural limitation. Specifically, the linear relationship between its inputs (temperature and net shortwave radiation) and output (melt) implied by its formulation breaks down with increasing cloud cover. In contrast, owing to a small difference in form, the simplified energy balance (SEB) method (Oerlemans 2001) retains a more stable relationship between the inputs and output under all cloud conditions. Albedo parameterisations aside, both the ETI and SEB formulations have two calibration parameters and the same input requirements. Thus the SEB approach could in theory be more robust. However, one of the parameters at least shows a dependence on cloud cover, which

would need to be accounted for, perhaps with a simple parameterisation, unless an "average" value for all cloud conditions were deemed acceptable.

For glacier ablation, the results in Chapter 5 show that the simple energy balance models for ablation of clean and debris-covered ice developed in this study offer melt responses that accord well with observed runoff. There is also general agreement with profiles of land surface temperature (LST) inferred from MODIS. However, similar to other studies (Collier et al. 2015), the vertical gradient of LST is not fully represented in the lower elevation glacier reaches. Similar to the case of snow, there is stronger agreement between the ETI model and energy balance approaches compared with the TI model, although considerable parameter equifinality is evident. Crucially, however, it is also shown that multiple ETI parameter combinations can result in relatively similar sensitivities to temperature perturbations. Moreover, these ETI sensitivities are much more similar to an energy balance formulation compared with the TI model. Although the simplicity of the climate scenarios tested must be acknowledged, this suggests that, despite its limitations, the ETI model likely represents a more suitable formulation for hydrological projections than the TI model. Nevertheless, a longer-term objective to produce hydrological projections with an energy balance model would seem worthwhile.

Objective 4: To analyse climate model skill in representing key controls on temperature variability and projections, focusing on the Karakoram Vortex

With spread in climate projections for the UIB being so wide, there is a need to understand how well the drivers of variability and trends are represented in climate models before assessing potential hydrological changes. This is a large, multi-dimensional problem. Addressing it means bringing together evaluations of how a vast array of processes and patterns are simulated in coupled land-atmosphere-ocean modelling. Of particular relevance to the UIB are model representations of the Karakoram Vortex (KV). The regional circulation anomalies manifest in the KV appear to be strongly implicated in the unique climatic, glaciological and hydrological trends discernible in the UIB in recent decades. Much is still to be investigated regarding the KV, but the underlying signature it embodies is gaining traction as an entry-point for understanding regionally differentiated climate patterns and trajectories (Mölg et al. 2017; Norris et al. 2018). Thus, assessing KV representation in climate models is a useful and necessary starting point for bringing analyses of critical processes shaping near-surface climate trajectories into hydrological projections.

Chapter 6 shows that the characteristic features of the KV are essentially reproduced in CMIP5 climate models, albeit to varying degrees. Crucially, the models capture the general patterns of correlation between a standardised zonal shear index, which quantifies KV states, and temperature anomalies throughout much of the troposphere. The associated anomaly patterns in horizontal and vertical winds are also present in the ensemble overall. This provides some initial indication that the anomalies in adiabatic processes connecting the KV and temperature responses are faithfully represented (Forsythe et al. 2017; Li et al. 2018). In addition, the models simulate the observed contraction from winter to summer in the area of positive correlation between KV circulation and temperature anomalies. Correctly reproducing this seasonal contraction is critical, as it is considered to be a substantial contributor to the divergent temperature trends across the Himalayan arc in recent decades (Forsythe et al. 2017).

However, there are some potential shortcomings in model performance. For the ensemble as a whole, the magnitude of correlation between KV and near-surface temperature anomalies is generally lower than in observations and reanalyses. Moreover, the spatial extent and centroid of correlation patterns characterising the KV vary notably between models. Although difficult to determine definitively, there is also a suggestion that simulated trends in temperature and the zonal shear index describing KV behaviour are less closely coupled in the CMIP5 models compared with observations. This is particularly the case in summer, when reanalyses and observations show cyclonic and stable/falling trends in KV behaviour and UIB temperature, respectively, over the second half of the twentieth century. In contrast, the CMIP5 models show warming in the absence of clear trends in KV behaviour. The very limited relationship between KV behaviour and temperature trends for models within the CMIP5 ensemble could indicate that near-surface temperature in the models shows heightened sensitivity to other factors in summer, as well as potentially insufficient sensitivity to KV variations. The results also raise other questions, such as whether the observed trends represent a response to ongoing climate change or a manifestation of natural variability, given that the circulation trends differ.

Analysing future projections indicates that the KV continues to modulate inter-annual variability in temperature in the CMIP5 models to a similar extent as in the historical period. However, other influences appear to be more important in shaping temperature trends. Focusing on summer as the season of anomalous recent historical temperature trends and runoff generation in the UIB, feedbacks between incoming shortwave radiation, albedo and

sensible heat fluxes are critical in projected trends. The results provide some suggestion that summer temperature change projections can be constrained with reference to representation of some historical land surface states and processes, particularly related to surface albedo. This is essentially in line with findings on the role of snow albedo feedbacks at larger scales, in terms of how inter-model variation offers the potential for applying observational constraints to projected changes (Hall and Qu 2006; Essery 2013; Qu and Hall 2014). This analysis suggests that similar methods may be applicable at the regional scale. In contrast to approaches taking cold/warm and dry/wet extremes to produce a very wide envelope of climate projections (e.g. Lutz et al. 2016a,b), there may thus be scope to reduce the range of possible climatic and hydrological futures projected for the Karakoram, UIB and neighbouring basins draining the Himalaya and Tibetan Plateau. This would be of substantial benefit for water resources planning, but resilience to a range of possible changes and variability will still need to be enhanced.

Objective 5: To explore the implications of selected global warming targets for the climate and hydrology of the UIB

Previous hydrological projections for the UIB have not explicitly considered changes in river flow regimes under the climate change targets set out in the Paris Agreement of 2015. More in line with current trajectories, we could also ask what the implications of 3 or 4°C warmer worlds may be. There is thus a need to frame the climate change impact problem in terms of stabilisation scenarios, with the ultimate aim of guiding both mitigation and adaptation efforts. This study provides the first analysis of climate and hydrological impacts in the UIB at 1.5°C, 2°C and higher global warming stabilisation targets using large ensembles of climate model simulations.

Chapter 7 indicates that climate models project amplified warming in the UIB compared with global temperature rises under the assessed stabilisation targets. Interestingly, the spread in simulated temperatures is much larger in winter/spring than in summer, according to the large HAPPI initial conditions ensembles. This may well relate to the strong controls on summer temperature changes exerted by model-specific snow and albedo climatology, variability and feedbacks discussed in Chapter 6. However, while temperature shows clear patterns of change in historical and future scenarios, precipitation does not show particularly clear shifts when the full HAPPI ensembles are considered. Thus, for CMIP5, where large initial conditions ensembles are unavailable for most participating models, it is unclear how much the projected

precipitation changes are simply reflections of natural variability, at least under the warming scenarios evaluated. Given the magnitude of spread in climatological precipitation in the HAPPI ensembles, it may well be that, over multi-decadal time scales, natural variability in precipitation will be a bigger challenge than any forced response to climate change.

Generally, the hydrological response to the warmer worlds evaluated is characterised by increasing flows in the earlier part of the melt season and decreasing flows in the later part. Note that this is under notionally stabilised climates; hydrological changes during the transitions from historical to future worlds are not explored in this project. Global warming scenarios of 1.5, 2 and 3°C lead to progressively more pronounced shifts in flow timing. When considering both sampling from the HAPPI initial conditions ensembles (to get a range of change factors) and the CMIP5 ensemble, the uncertainty in the mean response (and for drier/lower flow conditions) can become relatively large compared with the magnitude of inter-annual variability. With temperature perturbations alone, there are some notable groupings and coherence in the ensemble of hydrological responses, but the variability in response increases substantially when precipitation is taken into consideration. Building resilience to a range of potential shifts in flow timing will thus be necessary.

Based on CMIP5 and HAPPI ensemble means, there is not a large suggestion of notable changes in mean annual water availability. However, for drier years this begins to change under the 2 and 3°C warming scenarios, particularly for the Hunza sub-basin, where the buffering effect of glacier melt gets eroded by ice mass loss. More important perhaps is the variability in water availability in mean and drier conditions associated with natural climate variability. Exemplified by the CMIP5 ensemble, precipitation changes are the primary control on changes in annual water availability in the climate stabilisation scenarios. The increases in evapotranspiration in warmer worlds have less of an effect. There may be some relatively minor influence from the TOPKAPI-ETH evapotranspiration and soil moisture balance approach here, which could be investigated further.

Therefore, the fact that a large range of precipitation possibilities are projected by both CMIP5 and the HAPPI initial conditions ensembles suggest water management needs to account for the possibility of notable variability in mean annual flows over inter-annual to multi-decadal time scales. The overall shift in flow timing and the possibility of greater reductions in flow during drier periods are also particularly important water resources considerations. Given the challenges in meeting the needs of water users in the Indus under

present conditions, especially in drought periods (see Chapter 1), these findings suggest that developing new infrastructure to increase storage capacity in the basin is an urgent priority. This would help to buffer hydroclimatic variability and account for likely shifts in river flow timing over time. Whether resilience through enhanced storage can be achieved within the institutional context of Indus basin water management is a key question. If not, there is a strong possibility that severe shortfalls in water supply could occur in the coming decades during relatively dry periods, particularly given the ongoing rise in demand for water discussed in Chapter 1.

Transferability to Other Regions

A number of the findings presented here may be relevant for studies in other regions. Firstly, the delineation of the HAR's strengths and weaknesses may help to guide the design of similar WRF simulations for other mountainous contexts. This is particularly the case for regions with a pronounced annual cycle of snow cover, such as the European Alps or the North American and Canadian Rockies. As noted above, the results in this study suggest that improved snow state initialisation strategies and alternative land surface model choices could lead to better near-surface climate simulations. Secondly, the good performance of process-based snow modelling in the UIB opens up the possibility of taking a similar approach elsewhere. In particular, using a multi-physics ensemble and evaluating it with remote sensing of multiple land surface state variables may help to constrain process parameterisation choices and identify structural deficiencies in other data-sparse mountain regions. Thirdly, the findings on the behaviour and climate sensitivity of different model complexities may help to guide model selection for climate change impact studies. For example, the SEB variant of the ETI approach may ultimately be the most robust choice amongst the semi-empirical models, especially in cloudier contexts. Finally, the sensitivity of UIB hydrology to projected climate change even under low warming scenarios has implications for other catchments with cryosphere-dominated hydrological regimes in Asia and around the world. The results also indicate the importance of moving to the use of large ensembles of climate model simulations when conducting impact projections in mountainous contexts.

8.2 Further Work

The findings presented in this thesis raise a number of questions that could be investigated in further work. These are discussed in turn below.

Coupled Land-Atmosphere Modelling in the Himalaya

While the results presented suggest that the HAR exhibits many good features, it could potentially be improved in several ways. One of the most useful improvements might be revisions to the snow state initialisation approach, as well as snow process and sub-grid variability parameterisations. As discussed above, these aspects of the HAR methods and models appear to be linked to some of the key deficiencies in its performance. This concerns the seasonal variation in its bias patterns particularly. Initial steps to explore this might include examining how remote sensing could be used in initialisation of snow states and to parameterise snowpack processes. Indeed, this could draw on the findings regarding appropriate process representations derived from the offline FSM simulations in this study. Part of the solution may be switching from the Noah land surface model (LSM), whose snow process representation has admittedly improved over time (e.g. Livneh et al. 2010; Barlage et al. 2010; Wang et al. 2010), to the more recent Noah-MP or Community Land Model LSMs.

More generally, there is a need for further sensitivity testing of coupled land-atmosphere model inputs and physics parameterisation schemes in the UIB and surrounding region. These would ideally be conducted with a multi-physics approach to test different microphysics, cumulus, planetary boundary layer (PBL) and LSM schemes (e.g. García-Díez et al. 2015). It would also be interesting to examine further the role of boundary conditions and initialisation datasets (i.e. in addition to snow). This might include testing the recently released ERA5 reanalysis product, which could provide an improved forcing dataset. Together, these various avenues could form the basis for generating a multi-decadal, higher resolution successor to the HAR that uses improved driving data and physics schemes. Such a data product could be the biggest single way of spurring further advances in cryospheric and hydrological modelling in the UIB.

Cryospheric and Hydrological Modelling

There are several areas that could be investigated in attempts to improve further process-based snow modelling in the UIB. These include approaches to adjusting turbulent heat fluxes for atmospheric stability conditions. The most practical solution may be an intermediate adjustment that lies between the two end-point cases represented by the FSM options (e.g. Collier et al. 2015). While the data to parameterise this process directly do not really exist in this context, it could be that the multiple variables available from remote sensing provide a good enough basis for evaluation or calibration. However, testing of intermediate stability

adjustment options would ideally be preceded by further refinement of climate inputs. If some of the aforementioned issues with the HAR can be at least partly resolved, the additional confidence in model inputs may translate to additional confidence in the discrimination of snow process parameterisation.

Similarly, it is desirable for refined climate inputs to be developed before testing parameterisations of other processes. These include treatments of sub-grid variability, heat advection under conditions of patchy snow cover, terrain enhancement of longwave radiation, and deposition of aerosols and black carbon. They also include the effects of wind on snow redistribution and sublimation. While gravitational snow redistribution can be modelled fairly simply based primarily on terrain parameters, higher resolution simulations than the HAR would be better for incorporating wind effects. Incorporating these processes may help to better simulate the high elevation water balance, which is integral to the long-term trajectory of the basin's cryosphere. However, the complex nature of the climate in the UIB is such that model inputs will still be approximate even if HAR-type approaches can be refined. As such, if additional processes are incorporated it may be worth retaining an ensemble of roughly behavioural models to explore uncertainty in applications.

For glacier modelling, more attention should be devoted to accumulation and ice redistribution processes. While it has been shown that clean and debris-covered ice ablation can be modelled with formulations of varying complexity, alternative representations of snow-ice transitions could be explored. It would also be useful to investigate compare simple approaches to glacier mass redistribution, such as the Δh-parameterisation applied here (Huss et al. 2010), with solutions of simplified ice flow equations. In both cases, the emphasis could be more on whether significant divergences occur in climate change projections or other applications. This may be practically useful in a context where it is difficult to discriminate between model approaches on grounds of performance, given data limitations.

Climate and Hydrological Projections

For climate projections, it would be useful to develop further the climate model evaluations undertaken in the present work. It seems likely that model ensembles could be constrained at least to some degree using observations and process-based reasoning. For temperature projections, a key aspect of this is the representation of snow and the snow-albedo feedback. Improved observational constraints on the relevant processes may depend on further

developments in algorithms for retrieving surface albedo from remote sensing in complex terrain. This topic may be critical for refining climate projections in the UIB.

As understanding of the Karakoram Vortex and its relationship with regionally differentiated climate trajectories grows, this will also need to be investigated in climate models. In addition, it would be useful to draw on centennial reanalyses and paleoclimate records for the recent past to better characterise the nature of climate variability in the UIB and surrounding basins. This may provide some further contextualisation of recent climate changes, particularly in terms of their connection to anthropogenic warming compared with natural cycles. Characterisation of natural variability will be particularly important for managing the range of potential precipitation scenarios in the coming decades.

Further work should also seek to apply process-based snow and glacier models for hydrological projections, given their differing climate sensitivities compared with simpler models. Facilitating this depends on refined approaches to deriving climate scenarios, however. This could take the form of a weather generator or perhaps a dynamical or quasi-dynamical downscaling approach (Gutmann et al. 2016). Any method selected will need to take account of the complicated patterns of dependence and variability in climate fields in the complex terrain of the UIB. This would include accounting for elevation-dependent climate change (Pepin et al. 2015).

References

Abbott, M. B., J. C. Bathurst, J. A. Cunge, P. E. O'Connell, and J. Rasmussen, 1986: An introduction to the European Hydrological System—Systeme Hydrologique Europeen, "SHE", 1: History and philosophy of a physically-based, distributed modelling system. *J. Hydrol.*, **87**, 45–59, doi:10.1016/0022-1694(86)90114-9.

Akhter, M., 2017: The Political Ecology of the Water Scarcity/Security Nexus in the Indus Basin: Decentering Per Capita Water Supply. *Imagining Industan: Overcoming Water Insecurity in the Indus Basin*, Z. Adeel and R.G. Wirsing, Eds., Springer International Publishing, Switzerland, 21–33.

Ali, S., D. Li, F. Congbin, and F. Khan, 2015: Twenty first century climatic and hydrological changes over Upper Indus Basin of Himalayan region of Pakistan. *Environ. Res. Lett.*, **10**, 014007, doi:10.1088/1748-9326/10/1/014007.

Allen, R. G., L. S. Pereira, D. Raes, and M. Smith, 1998: *Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56*. FAO Rome 300(9), D05109 pp.

Anderson, K., and G. Peters, 2016: The trouble with negative emissions. *Science (80-.).*, **354**, 182–183, doi:10.1126/science.aah4567.

Andreadis, K. M., P. Storck, and D. P. Lettenmaier, 2009: Modeling snow accumulation and ablation processes in forested environments. *Water Resour. Res.*, **45**, W05429, doi:10.1029/2008WR007042.

Anwar, A. A., and M. T. Bhatti, 2018: Pakistan's Water Apportionment Accord of 1991: 25 Years and Beyond. *J. Water Resour. Plan. Manag.*, **144**, 05017015, doi:10.1061/(ASCE)WR.1943-5452.0000831.

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

Archer, D. R., 2004: Hydrological implications of spatial and altitudinal variation in temperature in the Upper Indus Basin. *Nord. Hydrol.*, **35**, 209–222, doi:10.2166/nh.2004.0015.

——, 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**, 47–61, doi:https://doi.org/10.5194/hess-8-47-2004.

Archer, D. R., and H. J. Fowler, 2008: Using meteorological data to forecast seasonal runoff on the River Jhelum, Pakistan. *J. Hydrol.*, **361**, 10–23, doi:10.1016/j.jhydrol.2008.07.017. http://dx.doi.org/10.1016/j.jhydrol.2008.07.017.

——, N. Forsythe, H. J. Fowler, and S. M. Shah, 2010: Sustainability of water resources

management in the Indus Basin under changing climatic and socio economic conditions. *Hydrol. Earth Syst. Sci.*, **14**, 1669–1680, doi:10.5194/hess-14-1669-2010.

Arendt, A., and Coauthors, 2015: *Randolph Glacier Inventory – A Dataset of Global Glacier Outlines: Version 5.0: Technical Report*. Colorado, USA,.

Arino, O., J. J. Ramos Perez, V. Kalogirou, S. Bontemps, P. Defourny, and E. Van Bogaert, 2012: Global Land Cover Map for 2009 (GlobCover 2009). doi:https://doi.org/10.1594/PANGAEA.787668. http://due.esrin.esa.int/page_globcover.php.

Ayala, A., F. Pellicciotti, S. MacDonell, J. McPhee, and P. Burlando, 2017: Patterns of glacier ablation across North-Central Chile: Identifying the limits of empirical melt models under sublimation-favourable conditions. *Water Resour. Res.*, **53**, 1–25, doi:10.1002/2016WR020126.

Azam, M. F., P. Wagnon, E. Berthier, C. Vincent, K. Fujita, and J. S. Kargel, 2018: Review of the status and mass changes of Himalayan- Karakoram glaciers. *J. Glaciol.*, **64**, 61–74, doi:10.1017/jog.2017.86.

Barlage, M., and Coauthors, 2010: Noah land surface model modifications to improve snowpack prediction in the Colorado Rocky Mountains. *J. Geophys. Res.*, **115**, D22101, doi:10.1029/2009JD013470.

Barlow, M., M. Wheeler, B. Lyon, and H. Cullen, 2005: Modulation of Daily Precipitation over Southwest Asia by the Madden-Julian Oscillation. *Mon. Weather Rev.*, **133**, 3579–3594, doi:10.1175/MWR3026.1.

Barnett, T. P., J. C. Adam, and D. P. Lettenmaier, 2005: Potential impacts of a warming climate on water availability in snow-dominated regions. *Nature*, **438**, 303–309, doi:10.1038/nature04141.

Bashir, F., X. Zeng, H. Gupta, and P. Hazenberg, 2017: A Hydrometeorological Perspective on the Karakoram Anomaly Using Unique Valley-Based Synoptic Weather Observations. *Geophys. Res. Lett.*, **44**, 10470–10478, doi:10.1002/2017GL075284.

Bastin, S., M. Chiriaco, and P. Drobinski, 2016: Control of radiation and evaporation on temperature variability in a WRF regional climate simulation: comparison with colocated long term ground based observations near Paris. *Clim. Dyn.*, 1–19, doi:10.1007/s00382-016-2974-1.

Benali, A., A. C. Carvalho, J. P. Nunes, N. Carvalhais, and A. Santos, 2012: Estimating air surface temperature in Portugal using MODIS LST data. *Remote Sens. Environ.*, **124**, 108–121, doi:10.1016/j.rse.2012.04.024. http://dx.doi.org/10.1016/j.rse.2012.04.024.

Bentsen, M., and Coauthors, 2013: The Norwegian Earth System Model, NorESM1-M – Part 1: Description and basic evaluation of the physical climate. *Geosci. Model Dev.*, **6**, 687–720, doi:10.5194/gmd-6-687-2013.

Bernhardt, M., and K. Schulz, 2010: SnowSlide: A simple routine for calculating gravitational snow transport. *Geophys. Res. Lett.*, **37**, L11502, doi:10.1029/2010GL043086.

——, ——, G. E. Liston, and G. Zängl, 2012: The influence of lateral snow redistribution processes on snow melt and sublimation in alpine regions. *J. Hydrol.*, **424-425**, 196–206, doi:10.1016/j.jhydrol.2012.01.001. http://dx.doi.org/10.1016/j.jhydrol.2012.01.001.

Betts, A. K., R. Desjardins, and D. Worth, 2013: Cloud radiative forcing of the diurnal cycle climate of the Canadian Prairies. *J. Geophys. Res. Atmos.*, **118**, 8935–8953, doi:10.1002/jgrd.50593.

——, ——, ——, S. Wang, and J. Li, 2014: Coupling of winter climate transitions to snow and clouds over the Prairies. *J. Geophys. Res. Atmos.*, **119**, 1118–1139, doi:10.1002/2013JD021168.

Beven, K., 1989: Changing Ideas in Hydrology - The Case of Physically-Based Models. *J. Hydrol.*, **105**, 157–172, doi:10.1016/0022-1694(89)90101-7.

——, 2001: How far can we go in distributed hydrological modelling? *Hydrol. Earth Syst. Sci.*, **5**, 1–12, doi:10.5194/hess-5-1-2001.

——, 2002: Towards an alternative blueprint for a physically based digitally simulated hydrologic response modelling system. *Hydrol. Process.*, **16**, 189–206, doi:10.1002/hyp.343.

Beven, K., 2006: Searching for the Holy Grail of scientific hydrology: Qt = H(S, R, dt)A as closure. *Hydrol. Earth Syst. Sci.*, **10**, 609–618, doi:10.5194/hess-10-609-2006.

Beven, K. J., 2012: *Rainfall-Runoff Modelling: The Primer*. 2nd ed. John Wiley and Sons, Chichester, United Kingdom,.

Beven, K. J., and M. J. Kirkby, 1979: A physically based, variable contributing area model of basin hydrology. *Hydrol. Sci. J.*, **24**, 43–69, doi:10.1080/02626667909491834.

Beven, K. J., and H. L. Cloke, 2012: Comment on "Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water"' by Eric F. Wood et al. *Water Resour. Res.*, **48**, W01801, doi:10.1029/2011WR010982.

Bhambri, R., T. Bolch, P. Kawishwar, D. P. Dobhal, D. Srivastava, and B. Pratap, 2013: Heterogeneity in glacier response in the upper Shyok valley, northeast Karakoram. *Cryosph.*, **7**, 1385–1398, doi:10.5194/tc-7-1385-2013.

Bhutiyani, M. R., V. S. Kale, and N. J. Pawar, 2010: Climate change and the precipitation variations in the northwestern Himalaya: 1866-2006. *Int. J. Climatol.*, **30**, 535–548, doi:10.1002/joc.1920.

Biskop, S., F. Maussion, P. Krause, and M. Fink, 2016: Differences in the water-balance components of four lakes in the southern-central Tibetan Plateau. *Hydrol. Earth Syst.*

Sci., **20**, 209–225, doi:10.5194/hess-20-209-2016.

Blandford, T. R., K. S. Humes, B. J. Harshburger, B. C. Moore, V. P. Walden, and H. Ye, 2008: Seasonal and Synoptic Variations in Near-Surface Air Temperature Lapse Rates in a Mountainous Basin. *J. Appl. Meteorol. Climatol.*, **47**, 249–261, doi:10.1175/2007JAMC1565.1.

Bocchiola, D., and G. Diolaiuti, 2013: Recent (1980-2009) evidence of climate change in the upper Karakoram, Pakistan. *Theor. Appl. Climatol.*, **113**, 611–641, doi:10.1007/s00704-012-0803-y.

Bocchiola, D., and Coauthors, 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**, 2059–2075, doi:10.5194/hess-15-2059-2011.

Boerst, U., 2013: Representation and Comparability of Climate Data collected in High Mountains. http://www.mountainresearchinitiative.org/en/perth/key-contact-workshops/2013-12-08-san-francisco-mri-kcw-prior-to-agu-fall-meeting/2386-boerst-uwe/file.

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

Bolch, T., T. Pieczonka, K. Mukherjee, and J. Shea, 2017: Brief communication: Glaciers in the Hunza catchment (Karakoram) have been nearly in balance since the 1970s. *Cryosph.*, **11**, 531–539, doi:10.5194/tc-11-531-2017.

Bollasina, M. A., Y. Ming, and V. Ramaswamy, 2011: Anthropogenic Aerosols and the Weakening of the South Asia Summer Monsoon. *Science (80-.).*, **334**, 502–506, doi:10.1126/science.1204994.

——, ——, and ——, 2013: Earlier onset of the Indian monsoon in the late twentieth century: The role of anthropogenic aerosols. *Geophys. Res. Lett.*, **40**, 3715–3720, doi:10.1002/grl.50719.

——, ——, V. Ramawamy, M. D. Schwarzkopf, and V. Naik, 2014: Contribution of local and remote anthropogenic aerosols to the twentieth century weakening of the South Asian Monsoon. *Geophys. Res. Lett.*, **41**, 680–687, doi:10.1002/2013GL058183.

Bookhagen, B., and D. W. Burbank, 2010: Toward a complete Himalayan hydrological budget: Spatiotemporal distribution of snowmelt and rainfall and their impact on river discharge. *J. Geophys. Res.*, **115**, F03019, doi:10.1029/2009JF001426.

Bothe, O., K. Fraedrich, and X. Zhu, 2011: Large-scale circulations and Tibetan Plateau summer drought and wetness in a high-resolution climate model. *Int. J. Climatol.*, **31**, 832–846, doi:10.1002/joc.2124.

Briscoe, J., 2010: Troubled Waters: Can a Bridge Be Built over the Indus? *Econ. Polit. Wkly.*, **45**, 28–32.

——, and U. Qamar, 2007: *Pakistan's Water Economy: Running Dry*. 1-155 pp.

Brock, B., M. Sharp, and I. C. Willis, 2000: Measurement and parameterization of albedo variations at Haut Glacier d'Arolla. *J. Glaciol.*, **46**, 675–688, doi:10.3189/172756500781832675.

Brock, B. W., I. C. Willis, and M. J. Sharp, 2006: Measurement and parameterization of aerodynamic roughness length variations at Haut Glacier d'Arolla, Switzerland. *J. Glaciol.*, **52**, 1–17, doi:10.3189/172756500781832675.

Brown, M. E., and Coauthors, 2014: An integrated modeling system for estimating glacier and snow melt driven streamflow from remote sensing and earth system data products in the Himalayas. *J. Hydrol.*, **519**, 1859–1869, doi:10.1016/j.jhydrol.2014.09.050. http://dx.doi.org/10.1016/j.jhydrol.2014.09.050.

Brun, F., E. Berthier, P. Wagnon, A. Kääb, and D. Treichler, 2017: A spatially resolved estimate of High Mountain Asia glacier mass balances from 2000 to 2016. *Nat. Geosci.*, **10**, 668–673, doi:10.1038/NGEO2999.

——, and Coauthors, 2018: Ice cliff contribution to the tongue-wide ablation of Changri Nup Glacier, Nepal, central Himalaya. *Cryosph.*, **12**, 3439–3457, doi:10.5194/tc-12-3439-2018.

Burlando, P., F. Pellicciotti, and U. Strasser, 2002: Modelling Mountainous Water Systems Between Learning and Speculating, Looking for Challenges. *Nord. Hydrol.*, **33**, 47–74, doi:10.2166/nh.2002.0004.

Cannon, F., L. M. V Carvalho, C. Jones, and B. Bookhagen, 2015: Multi-annual variations in winter westerly disturbance activity affecting the Himalaya. *Clim. Dyn.*, **44**, 441–455, doi:10.1007/s00382-014-2248-8.

——, ——, ——, A. Hoell, J. Norris, G. N. Kiladis, and A. A. Tahir, 2017a: The influence of tropical forcing on extreme winter precipitation in the western Himalaya. *Clim. Dyn.*, **48**, 1213–1232, doi:10.1007/s00382-016-3137-0. "http://dx.doi.org/10.1007/s00382-016-3137-0.

——, ——, ——, J. Norris, B. Bookhagen, and G. N. Kiladis, 2017b: Effects of topographic smoothing on the simulation of winter precipitation in High Mountain Asia. *J. Geophys. Res. Atmos.*, **122**, 1456–1474, doi:10.1002/2016JD026038.

Carenzo, M., F. Pellicciotti, S. Rimkus, and P. Burlando, 2009: Assessing the transferability and robustness of an enhanced temperature-index glacier-melt model. *J. Glaciol.*, **55**, 258–274, doi:10.3189/002214309788608804.

Carenzo, M., F. Pellicciotti, J. Mabillard, T. Reid, and B. W. Brock, 2016: An enhanced temperature index model for debris-covered glaciers accounting for thickness effect. *Adv. Water Resour.*, **94**, 457–469, doi:10.1016/j.advwatres.2016.05.001. http://dx.doi.org/10.1016/j.advwatres.2016.05.001.

Case, J. L., W. L. Crosson, S. V Kumar, W. M. Lapenta, and C. D. Peters-Lidard, 2008: Impacts of High-Resolution Land Surface Initialization on Regional Sensible Weather Forecasts from the WRF Model. *J. Hydrometeorol.*, **9**, 1249–1266, doi:10.1175/2008JHM990.1.

Ciarapica, L., and E. Todini, 2002: TOPKAPI: a model for the representation of the rainfall-runoff process at different scales. *Hydrol. Process.*, **16**, 207–229, doi:10.1002/hyp.342.

Clapp, R. B., and G. M. Hornberger, 1978: Empirical Equations for Some Soil Hydraulic Properties. *Water Resour. Res.*, **14**, 601–604, doi:10.1029/WR014i004p00601.

Clark, M. P., and Coauthors, 2011: Representing spatial variability of snow water equivalent in hydrologic and land-surface models: A review. *Water Resour. Res.*, **47**, W07539, doi:10.1029/2011WR010745.

——, and Coauthors, 2015a: A unified approach for process-based hydrologic modeling: 1. Modeling concept. *Water Resour. Res.*, **51**, 2498–2514, doi:10.1002/2015WR017200.A.

——, and Coauthors, 2015b: A unified approach for process-based hydrologic modeling: 2. Model implementation and case studies. *Water Resour. Res.*, **51**, 2515–2542, doi:10.1002/2015WR017198.A.

——, and Coauthors, 2016: Characterizing Uncertainty of the Hydrologic Impacts of Climate Change. *Curr. Clim. Chang. Reports*, **2**, 55–64, doi:10.1007/s40641-016-0034-x. http://dx.doi.org/10.1007/s40641-016-0034-x.

——, and Coauthors, 2017: The evolution of process-based hydrologic models: historical challenges and the collective quest for physical realism. *Hydrol. Earth Syst. Sci.*, **21**, 3427–3440, doi:10.5194/hess-21-3427-2017.

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

Collier, E., and W. W. Immerzeel, 2015: High-resolution modeling of atmospheric dynamics in the Nepalese Himalaya. *J. Geophys. Res. Atmos.*, **120**, 9882–9896, doi:10.1002/2015JD023266.

——, T. Mölg, F. Maussion, D. Scherer, C. Mayer, and A. B. G. Bush, 2013: High-resolution interactive modelling of the mountain glacier-atmosphere interface: an application over the Karakoram. *Cryosph.*, **7**, 779–795, doi:10.5194/tc-7-779-2013.

Collier, E., L. I. Nicholson, B. W. Brock, F. Maussion, R. Essery, and A. B. G. Bush, 2014: Representing moisture fluxes and phase changes in glacier debris cover using a reservoir approach. *Cryosph.*, **8**, 1429–1444, doi:10.5194/tc-8-1429-2014.

——, F. Maussion, L. I. Nicholson, T. Mölg, W. W. Immerzeel, and A. B. G. Bush, 2015: Impact of debris cover on glacier ablation and atmosphere-glacier feedbacks in the Karakoram. *Cryosph.*, **9**, 1617–1632, doi:10.5194/tc-9-1617-2015.

Condon, M., D. Kriens, A. Lohani, and E. Sattar, 2014: Challenge and response in the Indus Basin. *Water Policy*, **16**, 58–86, doi:10.2166/wp.2014.004.

Conway, J. P., and N. J. Cullen, 2013: Constraining turbulent heat flux parameterization over a temperate maritime glacier in New Zealand. *Ann. Glaciol.*, **54**, 41–51, doi:10.3189/2013AoG63A604.

Corripio, J. G., 2003: Vectorial algebra algorithms for calculating terrain parameters from DEMs and solar radiation modelling in mountainous terrain. *Int. J. Geogr. Inf. Sci.*, **17**, 1–23, doi:10.1080/713811744.

Cuffey, K., and W. S. B. Paterson, 2010a: *The Physics of Glaciers*. 4th Editio. Elsevier, Oxford, 704 pp.

Cuffey, K. M., and W. S. B. Paterson, 2010b: *The Physics of Glaciers*. Academic Press,.

Curio, J., and D. Scherer, 2016: Seasonality and spatial variability of dynamic precipitation controls on the Tibetan Plateau. *Earth Syst. Dyn.*, **7**, 767–782, doi:10.5194/esd-7-767-2016.

Curio, J., F. Maussion, and D. Scherer, 2015: A 12-year high-resolution climatology of atmospheric water transport over the Tibetan Plateau. *Earth Syst. Dyn.*, **6**, 109–124, doi:10.5194/esd-6-109-2015.

Dahri, Z. H., F. Ludwig, E. Moors, B. Ahmad, A. Khan, and P. Kabat, 2016: An appraisal of precipitation distribution in the high-altitude catchments of the Indus basin. *Sci. Total Environ.*, **548-549**, 289–306, doi:10.1016/j.scitotenv.2016.01.001. http://dx.doi.org/10.1016/j.scitotenv.2016.01.001.

——, E. Moors, F. Ludwig, S. Ahmad, A. Khan, I. Ali, and P. Kabat, 2018: Adjustment of measurement errors to reconcile precipitation distribution in the high-altitude Indus basin. *Int. J. Climatol.*, **38**, 3842–3860, doi:10.1002/joc.5539.

Daly, C., M. Halbleib, J. I. Smith, W. P. Gibson, M. K. Doggett, G. H. Taylor, J. Curtis, and P. P. Pasteris, 2008: Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. *Int. J. Climatol.*, **28**, 2031–2064, doi:10.1002/joc.1688.

——, M. E. Slater, J. A. Roberti, S. H. Laseter, and L. W. Swift Jr, 2017: High-resolution precipitation mapping in a mountainous watershed: ground truth for evaluating uncertainty in a national precipitation dataset. *Int. J. Climatol.*, **37**, 124–137, doi:10.1002/joc.4986.

Deb, K., A. Pratap, S. Agarwal, and T. Meyarivan, 2002: A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II. *IEEE Trans. Evol. Comput.*, **6**, 182–197, doi:10.1109/4235.996017.

Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.*, **137**, 553–597,

doi:10.1002/qj.828.

Dimri, A. P., D. Niyogi, A. P. Barros, J. Ridley, U. C. Mohanty, T. Yasunari, and D. R. Sikka, 2015: Western Disturbances: A review. *Rev. Geophys.*, **53**, 225–246, doi:10.1002/2014RG000460.

Ding, Q., and B. Wang, 2005: Circumglobal Teleconnection in the Northern Hemisphere Summer. *J. Clim.*, **18**, 3483–3505, doi:10.1175/JCLI3473.1.

Duethmann, D., J. Zimmer, A. Gafurov, A. Güntner, D. Kriegel, B. Merz, and S. Vorogushyn, 2013: Evaluation of areal precipitation estimates based on downscaled reanalysis and station data by hydrological modelling. *Hydrol. Earth Syst. Sci.*, **17**, 2415–2434, doi:10.5194/hess-17-2415-2013.

Essery, R., 2013: Large-scale simulations of snow albedo masking by forests. *Geophys. Res. Lett.*, **40**, 5521–5525, doi:10.1002/grl.51008.

——, 2015: A factorial snowpack model (FSM 1.0). *Geosci. Model Dev.*, **8**, 3867–3876, doi:10.5194/gmd-8-3867-2015.

——, and Coauthors, 2009: SNOWMIP2: An Evaluation of Forest Snow Process Simulations. *Bull. Am. Meteorol. Soc.*, **90**, 1120–1135, doi:https://doi.org/10.1175/2009BAMS2629.1.

——, S. Morin, Y. Lejeune, and C. B. Ménard, 2013: A comparison of 1701 snow models using observations from an alpine site. *Adv. Water Resour.*, **55**, 131–148, doi:10.1016/j.advwatres.2012.07.013. http://dx.doi.org/10.1016/j.advwatres.2012.07.013.

Etchevers, P., and Coauthors, 2004: Validation of the energy budget of an alpine snowpack simulated by several snow models (SnowMIP project). *Ann. Glaciol.*, **38**, 150–158, doi:https://doi.org/10.3189/172756404781814825.

Ewen, J., G. Parkin, and P. E. O'Connell, 2000: SHETRAN: Distributed River Basin Flow and Transport Modeling System. *J. Hydrol.*, **5**, 250–258, doi:10.1061/(ASCE)1084-0699(2000)5:3(250).

Farinotti, D., and Coauthors, 2017: How accurate are estimates of glacier ice thickness? Results from ITMIX, the Ice Thickness Models Intercomparison eXperiment. *Cryosph.*, **11**, 949–970, doi:10.5194/tc-11-949-2017.

Fatichi, S., and Coauthors, 2016: An overview of current applications, challenges, and future trends in distributed process-based models in hydrology. *J. Hydrol.*, **537**, 45–60, doi:10.1016/j.jhydrol.2016.03.026. http://dx.doi.org/10.1016/j.jhydrol.2016.03.026.

Figueres, C., H. J. Schellnhuber, G. Whiteman, J. Rockström, A. Hobley, and S. Rahmstorf, 2017: Three years to safeguard our climate. *Nature*, **546**, 593–595, doi:10.1038/546593a.

Filippi, L., E. Palazzi, J. von Hardenburg, and A. Provenzale, 2014: Multidecadal Variations in the Relationship between the NAO and Winter Precipitation in the Hindu Kush-Karakoram. *J. Clim.*, **27**, 7890–7902, doi:10.1175/JCLI-D-14-00286.1.

Finger, D., F. Pellicciotti, M. Konz, S. Rimkus, and P. Burlando, 2011: The value of glacier mass balance, satellite snow cover images, and hourly discharge for improving the performance of a physically based distributed hydrological model. *Water Resour. Res.*, **47**, W07519, doi:10.1029/2010WR009824.

Fischer, E. M., and R. Knutti, 2015: Anthropogenic contribution to global occurrence of heavy-precipitation and high-temperature extremes. *Nat. Clim. Chang.*, **5**, 560–565, doi:10.1038/NCLIMATE2617.

——, U. Beyerle, C. F. Schleussner, A. D. King, and R. Knutti, 2018: Biased Estimates of Changes in Climate Extremes From Prescribed SST Simulations. *Geophys. Res. Lett.*, **45**, 8500–8509, doi:10.1029/2018GL079176.

Fischer, G., F. Nachtergaele, S. Prieler, H. T. van Velthuizen, L. Verelst, and D. Wiberg, 2008: *Global Agro-ecological Zones Assessment for Agriculture (GAEZ 2008)*. IIASA, Laxenburg, Austria and FAO, Rome, Italy,.

Flato, G., and Coauthors, 2013: Evaluation of Climate Models. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, T.F. Stocker et al., Eds., Cambridge University Press, Cambridge, United Kingdom and New York, USA, 741–866.

Forsythe, N., H. J. Fowler, C. G. Kilsby, and D. R. Archer, 2012a: Opportunities from Remote Sensing for Supporting Water Resources Management in Village/Valley Scale Catchments in the Upper Indus Basin. *Water Resour. Manag.*, **26**, 845–871, doi:10.1007/s11269-011-9933-8.

——, C. G. Kilsby, H. J. Fowler, and D. R. Archer, 2012b: Assessment of Runoff Sensitivity in the Upper Indus Basin to Interannual Climate Variability and Potential Change Using MODIS Satellite Data Products. *Mt. Res. Dev.*, **32**, 16–29, doi:10.1659/MRD-JOURNAL-D-11-00027.1.

Forsythe, N., H. J. Fowler, S. Blenkinsop, A. Burton, C. G. Kilsby, D. R. Archer, C. Harpham, and M. Z. Hashmi, 2014: Application of a stochastic weather generator to assess climate change impacts in a semi-arid climate: The Upper Indus Basin. *J. Hydrol.*, **517**, 1019–1034, doi:10.1016/j.jhydrol.2014.06.031. http://dx.doi.org/10.1016/j.jhydrol.2014.06.031.

Forsythe, N., A. J. Hardy, H. J. Fowler, S. Blenkinsop, C. G. Kilsby, D. R. Archer, and M. Z. Hashmi, 2015: A Detailed Cloud Fraction Climatology of the Upper Indus Basin and Its Implications for Near-Surface Air Temperature. *J. Clim.*, **28**, 3537–3556, doi:10.1175/JCLI-D-14-00505.1.

——, 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**, 664–670, doi:10.1038/NCLIMATE3361.

Fowler, H. J., and D. R. Archer, 2005: Hydro-climatological variability in the Upper Indus Basin and implications for water resources. *Regional Hydrological Impacts of Climatic Change - Impact Assessment and Decision Making (Proceedings of symposium S6 held during the Seventh IAHS Scientific Assembly at Foz do Iguacu, Brazil, April 2005)*, Foz do Iguacu, Brazil, IAHS Publ. 295.

——, and ——, 2006: Conflicting Signals of Climatic Change in the Upper Indus Basin. *J. Clim.*, **19**, 4276–4293, doi:10.1175/JCLI3860.1.

Fowler, H. J., S. Blenkinsop, and C. Tebaldi, 2007: Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. *Int. J. Climatol.*, **27**, 1547–1578, doi:10.1002/joc.1556.

Gabbi, J., M. Carenzo, F. Pellicciotti, A. Bauder, and M. Funk, 2014: A comparison of empirical and physically based glacier surface melt models for long-term simulations of glacier response. *J. Glaciol.*, **60**, 1140–1154, doi:10.3189/2014JoG14J011.

Gafurov, A., and A. Bárdossy, 2009: Cloud removal methodology from MODIS snow cover product. *Hydrol. Earth Syst. Sci.*, **13**, 1361–1373, doi:10.5194/hess-13-1361-2009.

Gan, Y., and Coauthors, 2014: A comprehensive evaluation of various sensitivity analysis methods: A case study with a hydrological model. *Environ. Model. Softw.*, **51**, 269–285, doi:10.1016/j.envsoft.2013.09.031. http://dx.doi.org/10.1016/j.envsoft.2013.09.031.

García-Díez, M., J. Fernández, L. Fita, and C. Yagüe, 2013: Seasonal dependence of WRF model biases and sensitivity to PBL schemes over Europe. *Q. J. R. Meteorol. Soc.*, **139**, 501–514, doi:10.1002/qj.1976.

García-Díez, M., J. Fernández, and R. Vautard, 2015: An RCM multi-physics ensemble over Europe: multi-variable evaluation to avoid error compensation. *Clim. Dyn.*, **45**, 3141–3156, doi:10.1007/s00382-015-2529-x.

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

Gardelle, J., E. Berthier, Y. Arnaud, and A. Kääb, 2013: Region-wide glacier mass balances over the Pamir-Karakoram-Himalaya during 1999–2011. *Cryosph.*, **7**, 1263–1286, doi:10.5194/tc-7-1263-2013.

Gardner, A. S., and Coauthors, 2013: A Reconciled Estimate of Glacier Contributions to Sea Level Rise: 2003 to 2009. *Science (80-.).*, **340**, 852–858.

Gautam, R., N. C. Hsu, W. K. M. Lau, and T. J. Yasunari, 2013: Satellite observations of desert dust-induced Himalayan snow darkening. *Geophys. Res. Lett.*, **40**, 988–993, doi:10.1002/grl.50226.

Gelaro, R., and Coauthors, 2017: The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). *J. Clim.*, **30**, 5419–5454, doi:10.1175/JCLI-D-16-0758.1.

Gertler, C. G., S. P. Puppala, A. Panday, D. Stumm, and J. Shea, 2016: Black carbon and the Himalayan cryosphere: A review. *Atmos. Environ.*, **125**, 404–417, doi:10.1016/j.atmosenv.2015.08.078.

Ghatak, D., E. Sinsky, and J. Miller, 2014: Role of snow-albedo feedback in higher elevation warming over the Himalayas, Tibetan Plateau and Central Asia. *Environ. Res. Lett.*, **9**, 114008, doi:10.1088/1748-9326/9/11/114008.

Gibson, M. J., N. F. Glasser, D. J. Quincey, C. Mayer, A. V Rowan, and T. D. L. Irvine-fynn, 2017: Temporal variations in supraglacial debris distribution on Baltoro Glacier, Karakoram between 2001 and 2012. *Geomorphology*, **295**, 572–585, doi:10.1016/j.geomorph.2017.08.012. http://dx.doi.org/10.1016/j.geomorph.2017.08.012.

GOP, 2018: National Water Policy. 1–41.

Grayson, R. B., 1992: Physically Based Hydrologic Modeling 2. Is the Concept Realistic? *Water Resour. Res.*, **26**, 2659–2666, doi:10.1029/92WR01259.

Guerreiro, S. B., G. Kilsby, and F. Serinaldi, 2014: Analysis of time variation of rainfall in transnational basins in Iberia: abrupt changes or trends? *Int. J. Climatol.*, **34**, 114–133, doi:10.1002/joc.3669.

Guo, H., and Coauthors, 2018: Spring Snow-Albedo Feedback Analysis Over the Third Pole: Results From Satellite Observation and CMIP5 Model Simulations. *J. Geophys. Res. Atmos.*, **123**, 750–763, doi:10.1002/2017JD027846.

Gutmann, E., I. Barstad, M. Clark, J. Arnold, and R. Rasmussen, 2016: The Intermediate Complexity Atmospheric Research Model (ICAR). *J. Hydrometeorol.*, **17**, 957–973, doi:10.1175/JHM-D-15-0155.1.

Hall, A., and X. Qu, 2006: Using the current seasonal cycle to constrain snow albedo feedback in future climate change. *Geophys. Res. Lett.*, **33**, L03502, doi:10.1029/2005GL025127.

Hall, D. K., and G. A. Riggs, 2016: MODIS/Terra Snow Cover Daily L3 Global 500m Grid, Version 6. [2000-2015]. doi:10.5067/MODIS/MOD10A1.006.

Harder, P., J. W. Pomeroy, and W. Helgason, 2017: Local-Scale Advection of Sensible and Latent Heat During Snowmelt. *Geophys. Res. Lett.*, **44**, 9769–9777, doi:10.1002/2017GL074394.

Hasson, S., 2016a: Seasonality of Precipitation over Himalayan Watersheds in CORDEX South Asia and their Driving. *Atmosphere (Basel).*, **7**, 1–29, doi:10.3390/atmos7100123.

——, 2016b: Future Water Availability from Hindukush-Karakoram-Himalaya upper Indus Basin under Conflicting Climate Change Scenarios. *Climate*, **4**, 1–28, doi:10.3390/cli4030040.

——, J. Böhner, and F. Chishtie, 2018: Low fidelity of CORDEX and their driving experiments indicates future climatic uncertainty over Himalayan watersheds of Indus basin. *Clim. Dyn.*, 1–22, doi:10.1007/s00382-018-4160-0. http://dx.doi.org/10.1007/s00382-018-4160-0.

Hasson, S. ul, S. Pascale, V. Lucarini, and J. Böhner, 2016: Seasonal cycle of precipitation over major river basins in South and Southeast Asia: A review of the CMIP5 climate models data for present climate and future climate projections. *Atmos. Res.*, **180**, 42–63, doi:10.1016/j.atmosres.2016.05.008. http://dx.doi.org/10.1016/j.atmosres.2016.05.008.

——, J. Böhner, and V. Lucarini, 2017: Prevailing climatic trends and runoff response from Hindukush–Karakoram–Himalaya, upper Indus basin. *Earth Syst. Dyn.*, **8**, 337–355, doi:10.5194/esd-8-337-2017. http://www.earth-syst-dynam.net/8/337/2017/.

Hewitt, K., 1982: Natural dams and outburst floods of the Karakoram Himalaya. *Hydrological Aspects of Alpine and High Mountain Areas (Proceedings of the Exeter Symposium, July 1982)*, 259–269.

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

——, 2011: Glacier Change, Concentration, and Elevation Effects in the Karakoram Himalaya, Upper Indus Basin. *Mt. Res. Dev.*, **31**, 188–200, doi:10.1659/MRD-JOURNAL-D-11-00020.1.

——, 2014: *Glaciers of the Karakoram Himalaya*. Springer Netherlands,.

Heynen, M., E. Miles, S. Ragettli, P. Buri, W. W. Immerzeel, and F. Pellicciotti, 2016: Air temperature variability in a high-elevation Himalayan catchment. *Ann. Glaciol.*, **57**, 212–222, doi:10.3189/2016AoG71A076.

Hock, R., 2003: Temperature index melt modelling in mountain areas. *J. Hydrol.*, **282**, 104–115, doi:10.1016/S0022-1694(03)00257-9.

Hrachowitz, M., and M. P. Clark, 2017: HESS Opinions: The complementary merits of competing modelling philosophies in hydrology. *Hydrol. Earth Syst. Sci.*, **21**, 3953–3973, doi:10.5194/hess-21-3953-2017.

Hu, X., J. W. Nielsen-Gammon, and F. Zhang, 2010: Evaluation of Three Planetary Boundary Layer Schemes in the WRF Model. *J. Appl. Meteorol. Climatol.*, **49**, 1831–1844, doi:10.1175/2010JAMC2432.1.

Huffman, G. J., and Coauthors, 2007: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine

Scales. *J. Hydrometeorol.*, **8**, 38–55, doi:10.1175/JHM560.1.

Huintjes, E., and Coauthors, 2015: Evaluation of a coupled snow and energy balance model for Zhadang glacier, Tibetan Plateau, using glaciological measurements and time-lapse photography. *Arctic, Antarct. Alp. Res.*, **47**, 573–590, doi:10.1657/AAAR0014-073.

Hulme, M., 2016: 1.5°C and climate research after the Paris Agreement. *Nat. Clim. Chang.*, **6**, 222–224, doi:10.1038/nclimate2939. http://dx.doi.org/10.1038/nclimate2939.

Hunt, K. M. R., A. G. Turner, and L. C. Shaffrey, 2018a: Extreme Daily Rainfall in Pakistan and North India: Scale Interactions, Mechanisms, and Precursors. *Mon. Weather Rev.*, **146**, 1005–1022, doi:10.1175/MWR-D-17-0258.1.

——, G. Turner, and L. C. Shaffrey, 2018b: The evolution, seasonality and impacts of western disturbances. *Q. J. R. Meteorol. Soc.*, **144**, 278–290, doi:10.1002/qj.3200.

Huss, M., G. Jouvet, D. Farinotti, and A. Bauder, 2010: Future high-mountain hydrology: a new parameterization of glacier retreat. *Hydrol. Earth Syst. Sci.*, **14**, 815–829, doi:10.5194/hess-14-815-2010.

Ikeda, K., and Coauthors, 2010: Simulation of seasonal snowfall over Colorado. *Atmos. Res.*, **97**, 462–477, doi:10.1016/j.atmosres.2010.04.010.

Immerzeel, W. W., P. Droogers, S. M. de Jong, and M. F. P. Bierkens, 2009: Large-scale monitoring of snow cover and runoff simulation in Himalayan river basins using remote sensing. *Remote Sens. Environ.*, **113**, 40–49, doi:10.1016/j.rse.2008.08.010. http://dx.doi.org/10.1016/j.rse.2008.08.010.

Immerzeel, W. W., L. P. H. van Beek, and M. F. P. Bierkens, 2010: Climate Change Will Affect the Asian Water Towers. *Science (80-.).*, **382**, 1382–1385, doi:10.1126/science.1183188.

——, L. P. H. Van Beek, M. Konz, A. B. Shrestha, and M. F. P. Bierkens, 2012a: Hydrological response to climate change in a glacierized catchment in the Himalayas. *Clim. Change*, **110**, 721–736, doi:10.1007/s10584-011-0143-4.

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

Immerzeel, W. W., F. Pellicciotti, and M. F. P. Bierkens, 2013: Rising river flows throughout the twenty-first century in two Himalayan glacierized watersheds. *Nat. Geosci.*, **6**, 742–745, doi:10.1038/ngeo1896. http://dx.doi.org/10.1038/ngeo1896.

Immerzeel, W. W., L. Petersen, S. Ragettli, and F. Pellicciotti, 2014: The importance of observed gradients of air temperature and precipitation for modeling runoff from a glacierized watershed in the Nepalese Himalayas. *Water Resour. Res.*, **50**, 2212–2226, doi:10.1002/2013WR014506.

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**, 4673–4687, doi:10.5194/hess-19-4673-2015.

Inam, A., P. D. Clift, L. Giosan, A. R. Tabrez, M. Tahir, M. M. Rabbani, and M. Danish, 2007: The Geographic, Geological and Oceanographic Setting of the Indus River. *Large Rivers: Geomorphology and Management*, A. Gupta, Ed., John Wiley and Sons, 333–346.

IPCC, 2013: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. T.F. Stocker et al., Eds. Cambridge University Press, Cambridge, United Kingdom and New York, USA, 1535 pp.

——, 2018: *Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change,*. V. Masson-Delmotte et al., Eds.

James, R., R. Washington, C.-F. Schleussner, J. Rogelj, and D. Conway, 2017: Characterizing half-a-degree difference: a review of methods for identifying regional climate responses to global warming targets. *WIREs Clim. Chang.*, **8**, 457, doi:10.1002/wcc.457.

Jarosch, A. H., C. G. Schoof, and F. S. Anslow, 2013: Restoring mass conservation to shallow ice flow models over complex terrain. *Cryosph.*, **7**, 229–240, doi:10.5194/tc-7-229-2013.

Jarvis, A., H. I. Reuter, A. Nelson, and E. Guevara, 2008: Hole-filled SRTM for the globe Version 4, available from the CGIAR-CSI SRTM 90m Database (http://srtm.csi.cgiar.org).

Jordan, R., 1991: *A one-dimensional temperature model for a snow cover: Technical documentation for SNTHERM89, Tech. Rep. 91-16*. Hanover, N. H.,.

Kääb, A., E. Berthier, C. Nuth, J. Gardelle, and Y. Arnaud, 2012: Contrasting patterns of early twenty-first-century glacier mass change in the Himalayas. *Nature*, **488**, 495–498, doi:10.1038/nature11324.

Kääb, A., D. Treichler, C. Nuth, and E. Berthier, 2015: Brief Communication: Contending estimates of 2003–2008 glacier mass balance over the Pamir–Karakoram–Himalaya. *Cryosph.*, **9**, 557–564, doi:10.5194/tc-9-557-2015.

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 the Himalayas due to a unique seasonal cycle. *Nat. Geosci.*, **7**, 834–840, doi:10.1038/NGEO2269.

Karki, M. B., A. B. Shrestha, and M. Winiger, 2011: Enhancing Knowledge Management and Adaptation Capacity for Integrated Management of Water Resources in the Indus River Basin. *Mt. Res. Dev.*, **31**, 242–251, doi:10.1659/MRD-JOURNAL-D-11-00017.1.

Karki, R., S. Hasson, L. Gerlitz, U. Schickhoff, T. Scholten, and J. Böhner, 2017: Quantifying the added value of convection-permitting climate simulations in complex terrain: a systematic evaluation of WRF over the Himalayas. *Earth Syst. Dyn.*, **8**, 507–528, doi:10.5194/esd-8-507-2017.

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.

——, B. S. Naz, and L. C. Bowling, 2015: Separating snow, clean and debris covered ice in the Upper Indus Basin, Hindukush-Karakoram-Himalayas, using Landsat images between 1998 and 2002. *J. Hydrol.*, **521**, 46–64, doi:10.1016/j.jhydrol.2014.11.048. http://dx.doi.org/10.1016/j.jhydrol.2014.11.048.

Khattak, M. S., M. S. Babel, and M. Sharif, 2011: Hydro-meteorological trends in the upper Indus River basin in Pakistan. *Clim. Res.*, **46**, 103–119, doi:10.3354/cr00957.

Kilsby, C. G., and Coauthors, 2007: A daily weather generator for use in climate change studies. *2Environmental Model. Softw.*, **22**, 1705–1719, doi:10.1016/j.envsoft.2007.02.005.

King, A. D., D. J. Karoly, and B. J. Henley, 2017: Australian climate extremes at 1.5°C and 2°C of global warming. *Nat. Clim. Chang.*, **7**, 412–416, doi:10.1038/NCLIMATE3296.

Klemeš, V., 1990: The modelling of mountain hydrology: the ultimate challenge. *Hydrol. Mt. Areas (Proceedings Štrbské Pleso Work. Czechoslov. June 1989)*, **IAHS 190**, 29–44.

Kobayashi, S., and Coauthors, 2015: The JRA-55 Reanalysis: General Specifications and Basic Characteristics. *J. Meteorol. Soc. Japan*, **93**, 5–48, doi:10.2151/jmsj.2015-001.

Kraaijenbrink, P. D. A., M. F. P. Bierkens, A. F. Lutz, and W. W. Immerzeel, 2017: Impact of a global temperature rise of 1.5 degrees Celsius on Asia's glaciers. *Nature*, **549**, 257–260, doi:10.1038/nature23878. http://dx.doi.org/10.1038/nature23878.

Krishnan, R., V. Kumar, M. Sugi, and J. Yoshimura, 2009: Internal Feedbacks from Monsoon-Midlatitude Interactions during Droughts in the Indian Summer Monsoon. *J. Atmos. Sci.*, **66**, 553–578, doi:10.1175/2008JAS2723.1.

Lafaysse, M., B. Cluzet, M. Dumont, Y. Lejeune, V. Vionnet, and S. Morin, 2017: A multiphysical ensemble system of numerical snow modelling. *Cryosph.*, **11**, 1173–1198, doi:10.5194/tc-11-1173-2017.

Lang, H., and L. Braun, 1990: On the information content of air temperature in the context of snow melt estimation. *Hydrology of Mountainous Areas (Proceedings of the Štrbské Pleso Workshop, Czechoslovakia, June 1988)*, IAHS Publ. 190, 347–354.

Lang, T. J., and A. P. Barros, 2004: Winter Storms in the Central Himalayas. *J. Meteorol. Soc. Japan*, **82**, 829–844, doi:10.2151/jmsj.2004.829.

Lee, D., S. Min, E. Fischer, H. Shiogama, I. Bethke, L. Lierhammer, and J. F. Scinocca, 2018: Impacts of half a degree additional warming on the Asian summer monsoon rainfall characteristics. *Environ. Res. Lett.*, **13**, 044033, doi:10.1088/1748-9326/aab55d.

Lee, J.-Y., B. Wang, K.-H. Seo, J.-S. Kug, Y.-S. Choi, Y. Kosaka, and K.-J. Ha, 2014: Future Change of Northern Hemisphere Summer Tropical-Extratropical Teleconnection in CMIP5 Models. *J. Clim.*, **27**, 3643–3664, doi:10.1175/JCLI-D-13-00261.1.

Lehning, M., V. Ingo, D. Gustafsson, T. A. Nguyen, M. Stähli, and M. Zappa, 2006: ALPINE3D: a detailed model of mountain surface processes and its application to snow hydrology. *Hydrol. Process.*, **20**, 2111–2128, doi:10.1002/hyp.6204.

Levine, R. C., A. G. Turner, D. Marathayil, and G. M. Martin, 2013: The role of northern Arabian Sea surface temperature biases in CMIP5 model simulations and future projections of Indian summer monsoon rainfall. *Clim. Dyn.*, **41**, 155–172, doi:10.1007/s00382-012-1656-x.

Li, X.-F., H. Fowler, N. Forsythe, S. Blenkinsop, and D. Pritchard, 2018: The Karakoram/Western Tibetan vortex: seasonal and year-to-year variability. *Clim. Dyn.*, 1–24, doi:10.1007/s00382-018-4118-2.

Li, Z., and Coauthors, 2016: Aerosol and monsoon climate interactions over Asia. *Rev. Geophys.*, **54**, 866–929, doi:10.1002/2015RG000500.

Liston, G. E., and K. Elder, 2006a: A Distributed Snow-Evolution Modeling System (SnowModel). *J. Hydrometeorol.*, **7**, 1259–1276, doi:10.1175/JHM548.1.

——, and K. Elder, 2006b: A Meteorological Distribution System for High-Resolution Terrestrial Modeling (MicroMet). *J. Hydrometeorol.*, **7**, 217–234, doi:10.1175/JHM486.1.

Liu, J., and Coauthors, 2009: Validation of Moderate Resolution Imaging Spectroradiometer (MODIS) albedo retrieval algorithm: Dependence of albedo on solar zenith angle. *J. Geophys. Res.*, **114**, D01106, doi:10.1029/2008JD009969.

Liu, Z., and E. Todini, 2002: Towards a comprehensive physically-based rainfall-runoff model. *Hydrol. Earth Syst. Sci.*, **6**, 859–881, doi:10.5194/hess-6-859-2002.

——, M. L. V Martina, and E. Todini, 2005: Flood forecasting using a fully distributed model: application of the TOPKAPI model to the Upper Xixian Catchment. *Hydrol. Earth Syst. Sci.*, **9**, 347–364, doi:10.5194/hess-9-347-2005.

Livneh, B., Y. Xia, K. E. Mitchell, M. B. Ek, and D. P. Lettenmaier, 2010: Noah LSM Snow Model Diagnostics and Enhancements. *J. Hydrometeorol.*, **11**, 721–738, doi:10.1175/2009JHM1174.1.

López-Moreno, J. I., and Coauthors, 2017: Different sensitivities of snowpacks to warming in Mediterranean climate mountain areas. *Environ. Res. Lett.*, **12**, 074006, doi:10.1088/1748-9326/aa70cb.

Lute, A. C., and C. H. Luce, 2017: Are Model Transferability and Complexity Antithetical? Insights from Validation of a Variable-Complexity Empirical Snow Model in Space and Time. *Water Resour. Res.*, **53**, 8825–8850, doi:10.1002/2017WR020752.

Lutz, A. F., W. W. Immerzeel, A. B. Shrestha, and M. F. P. Bierkens, 2014: Consistent increase in High Asia's runoff due to increasing glacier melt and precipitation. *Nat. Clim. Chang.*, **4**, 587–592, doi:10.1038/NCLIMATE2237.

Lutz, A. F., W. W. Immerzeel, P. D. A. Kraaijenbrink, A. B. Shrestha, and M. F. P. Bierkens, 2016a: Climate Change Impacts on the Upper Indus Hydrology: Sources, Shifts and Extremes. *PLoS One*, **11**, 1–33, doi:10.1371/journal.pone.0165630.

——, H. W. ter Maat, H. Biemans, A. B. Shrestha, P. Wester, and W. W. Immerzeel, 2016b: Selecting representative climate models for climate change impact studies: an advanced envelope-based selection approach. *Int. J. Climatol.*, **36**, 3988–4005, doi:10.1002/joc.4608.

Macdonald, M. K., J. W. Pomeroy, and A. Pietroniro, 2010: On the importance of sublimation to an alpine snow mass balance in the Canadian Rocky Mountains. *Hydrol. Earth Syst. Sci.*, **14**, 1401–1415, doi:10.5194/hess-14-1401-2010.

MacDougall, A. H., and G. E. Flowers, 2010: Spatial and Temporal Transferability of a Distributed Energy-Balance Glacier Melt Model. *J. Clim.*, **24**, 1480–1499, doi:10.1175/2010JCLI3821.1.

MacDougall, A. H., B. A. Wheler, and G. E. Flowers, 2011: A preliminary assessment of glacier melt-model parameter sensitivity and transferability in a dry subarctic environment. *Cryosph.*, **5**, 1011–1028, doi:10.5194/tc-5-1011-2011.

Madhura, R. K., R. Krishnan, J. V Revadekar, M. Mujumdar, and B. N. Goswami, 2015: Changes in western disturbances over the Western Himalayas in a warming environment. *Clim. Dyn.*, **44**, 1157–1168, doi:10.1007/s00382-014-2166-9.

Magnusson, J., N. Wever, R. Essery, N. Helbig, A. Winstral, and T. Jonas, 2015: Evaluating snow models with varying process representations for hydrological applications. *Water Resour. Res.*, **51**, 2707–2723, doi:10.1002/2014WR016498.

Maidment, D. R., 1993: *Handbook of Hydrology*. McGraw-Hill, New York, 1424 pp.

Maraun, D., and Coauthors, 2010: Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user. *Rev. Geophys.*, **48**, RG3003, doi:10.1029/2009RG000314.

Massey, J. D., S. W. J, J. C. Knievel, and Y. Y. Cheng, 2016: Regional Soil Moisture Biases and Their Influence on WRF Model Temperature Forecasts over the Intermountain West. *Weather Forecast.*, **31**, 197–216, doi:10.1175/WAF-D-15-0073.1.

Matthews, T., R. Hodgkins, R. L. Wilby, S. Guðmundsson, F. Pálsson, H. Björnsson, and S. Carr, 2015: Conditioning temperature-index model parameters on synoptic weather types

for glacier melt simulations. *Hydrol. Process.*, **29**, 1027–1045, doi:10.1002/hyp.10217.

Mauritsen, T., and R. Pincus, 2017: Committed warming inferred from observations. *Nat. Clim. Chang.*, doi:10.1038/NCLIMATE3357.

Maussion, F., D. Scherer, R. Finkelnburg, J. Richters, W. Yang, and T. Yao, 2011: WRF simulation of a precipitation event over the Tibetan Plateau, China – an assessment using remote sensing and ground observations. *Hydrol. Earth Syst. Sci.*, **15**, 1795–1817, doi:10.5194/hess-15-1795-2011.

Maussion, F., D. Scherer, T. Mölg, E. Collier, J. Curio, and R. Finkelnburg, 2014: Precipitation Seasonality and Variability over the Tibetan Plateau as Resolved by the High Asia Reanalysis. *J. Clim.*, **27**, 1910–1927, doi:10.1175/JCLI-D-13-00282.1.

Meng, X., and Coauthors, 2018: Simulated cold bias being improved by using MODIS time-varying albedo in the Tibetan Plateau in WRF model. *Environ. Res. Lett.*, **13**, 044028, doi:10.1088/1748-9326/aab44a.

Mihalcea, C., C. Mayer, G. Diolaiuti, A. Lambrecht, C. Smiraglia, and G. Tartari, 2006: Ice ablation and meteorological conditions on the debris-covered area of Baltoro glacier, Karakoram, Pakistan. *Ann. Glaciol.*, **43**, 292–300, doi:10.3189/172756406781812104.

Mihalcea, C., C. Mayer, G. Diolaiuti, C. D'Agata, C. Smiraglia, A. Lambrecht, E. Vuillermoz, and G. Tartari, 2008: Spatial distribution of debris thickness and melting from remote-sensing and meteorological data, at debris-covered Baltoro glacier, Karakoram, Pakistan. *Ann. Glaciol.*, **48**, 49–57, doi:10.3189/172756408784700680.

Miles, E. S., F. Pellicciotti, I. C. Willis, J. F. Steiner, P. Buri, and N. S. Arnold, 2016: Refined energy-balance modelling of a supraglacial pond, Langtang Khola, Nepal. *Ann. Glaciol.*, **57**, 29–40, doi:10.3189/2016AoG71A421.

Minora, U., and Coauthors, 2015: A simple model to evaluate ice melt over the ablation area of glaciers in the Central Karakoram National Park, Pakistan. *Ann. Glaciol.*, **56**, 202–216, doi:10.3189/2015AoG70A206.

Mishra, V., 2015: Climatic uncertainty in Himalayan water towers. *J. Geophys. Res. Atmos.*, **120**, 2689–2705, doi:10.1002/2014JD022650.

Mitchell, D., R. James, P. M. Forster, R. A. Betts, H. Shiogama, and M. Allen, 2016: Realizing the impacts of a 1.5°C warmer world. *Nat. Clim. Chang.*, **6**, 735–737, doi:10.1038/nclimate3055. http://dx.doi.org/10.1038/nclimate3055.

——, and Coauthors, 2017: Half a degree additional warming, prognosis and projected impacts (HAPPI): background and experimental design. *Geosci. Model Dev.*, **10**, 571–583, doi:10.5194/gmd-10-571-2017.

Moeser, D., G. Mazzotti, N. Helbig, and T. Jonas, 2016: Representing spatial variability of forest snow: Implementation of a new interception model. *Water Resour. Res.*, **52**, 1208–1226, doi:10.1002/2015WR017961.

Mölg, T., and G. Kaser, 2011: A new approach to resolving climate-cryosphere relations: Downscaling climate dynamics to glacier-scale mass and energy balance without statistical scale linking. *J. Geophys. Res.*, **116**, D16101, doi:10.1029/2011JD015669.

——, N. J. Cullen, D. R. Hardy, G. Kaser, and L. Klok, 2008: Mass balance of a slope glacier on Kilimanjaro and its sensitivity to climate. *Int. J. Climatol.*, **28**, 881–892, doi:10.1002/joc.1589.

——, ——, ——, M. Winkler, and G. Kaser, 2009: Quantifying Climate Change in the Tropical Midtroposphere over East Africa from Glacier Shrinkage on Kilimanjaro. *J. Clim.*, **22**, 4162–4181, doi:10.1175/2009JCLI2954.1.

——, F. Maussion, and D. Scherer, 2014: Mid-latitude westerlies as a driver of glacier variability in monsoonal High Asia. *Nat. Clim. Chang.*, **4**, 68–73, doi:10.1038/nclimate2055.

——, ——, E. Collier, J. C. H. Chiang, and D. Scherer, 2017: Prominent Midlatitude Circulation Signature in High Asia's Surface Climate During Monsoon. *J. Geophys. Res. Atmos.*, **122**, 12702–12712, doi:10.1002/2017JD027414.

Montanari, A., and D. Koutsoyiannis, 2012: A blueprint for process-based modeling of uncertain hydrological systems. *Water Resour. Res.*, **48**, W09555, doi:10.1029/2011WR011412.

Morris, M. D., 1991: Factorial Sampling Plans for Preliminary Computational Experiments. *Technometrics*, **33**, 161–174, doi:10.1080/00401706.1991.10484804.

Mukhopadhyay, B., and A. Khan, 2014: Rising river flows and glacial mass balance in central Karakoram. *J. Hydrol.*, **513**, 192–203, doi:10.1016/j.jhydrol.2014.03.042. http://dx.doi.org/10.1016/j.jhydrol.2014.03.042.

——, and ——, 2015: A reevaluation of the snowmelt and glacial melt in river flows within Upper Indus Basin and its significance in a changing climate. *J. Hydrol.*, **527**, 119–132, doi:10.1016/j.jhydrol.2015.04.045. http://dx.doi.org/10.1016/j.jhydrol.2015.04.045.

——, and ——, 2016: Altitudinal variations of temperature, equilibrium line altitude, and accumulation-area ratio in Upper Indus Basin. *Hydrol. Res.*, **48**, 214–230, doi:10.2166/nh.2016.144.

——, ——, and R. Gautam, 2015: Rising and falling river flows: contrasting signals of climate change and glacier mass balance from the eastern and western Karakoram. *Hydrol. Sci. J.*, **60**, 2062–2085, doi:http://dx.doi.org/10.1080/02626667.2014.947291.

Neale, R. B., J. Richter, S. Park, P. H. Lauritzen, S. J. Vavrus, P. J. Rasch, and M. Zhang, 2013: The Mean Climate of the Community Atmosphere Model (CAM4) in Forced SST and Fully Coupled Experiments. *J. Clim.*, **26**, 5150–5168, doi:10.1175/JCLI-D-12-00236.1.

NOAA/NCEP, 2000: NCEP FNL Operational Model Global Tropospheric Analyses,

continuing from July 1999 (updated daily). doi:10.5065/D6M043C6.

Norris, J., L. M. V Carvalho, C. Jones, and F. Cannon, 2015: WRF simulations of two extreme snowfall events associated with contrasting extratropical cyclones over the western and central Himalaya. *J. Geophys. Res. Atmos.*, **120**, 3114–3138, doi:10.1002/2014JD022592.

Norris, J., L. M. V Carvalho, C. Jones, F. Cannon, B. Bookhagen, E. Palazzi, and A. A. Tahir, 2016: The spatiotemporal variability of precipitation over the Himalaya: evaluation of one-year WRF model simulation. *Clim. Dyn.*, **49**, 2179–2204, doi:10.1007/s00382-016-3414-y.

——, ——, ——, and ——, 2018: Deciphering the contrasting climatic trends between the central Himalaya and Karakoram with 36 years of WRF simulations. *Clim. Dyn.*, 1–22, doi:10.1007/s00382-018-4133-3. http://dx.doi.org/10.1007/s00382-018-4133-3.

Oerlemans, J., 2001: *Glaciers and climate change*. AA Balkema, Lisse,.

Otkin, J. A., and T. J. Greenwald, 2008: Comparison of WRF Model-Simulated and MODIS-Derived Cloud Data. *Mon. Weather Rev.*, **136**, 1957–1970, doi:10.1175/2007MWR2293.1.

Palazzi, E., J. Von Hardenberg, and A. Provenzale, 2013: Precipitation in the Hindu-Kush Karakoram Himalaya: Observations and future scenarios. *J. Geophys. Res. Atmos.*, **118**, 85–100, doi:10.1029/2012JD018697.

Palazzi, E., J. Von Hardenberg, S. Terzago, and A. Provenzale, 2015: Precipitation in the Karakoram-Himalaya: a CMIP5 view. *Clim. Dyn.*, **45**, 21–45, doi:10.1007/s00382-014-2341-z.

——, L. Filippi, and J. Von Hardenberg, 2017: Insights into elevation-dependent warming in the Tibetan Plateau-Himalayas from CMIP5 model simulations. *Clim. Dyn.*, **48**, 3991–4008, doi:10.1007/s00382-016-3316-z. "http://dx.doi.org/10.1007/s00382-016-3316-z.

Paltan, H., M. Allen, K. Haustein, L. Fuldauer, and S. Dadson, 2018: Global implications of 1.5°C and 2°C warmer worlds on extreme river flows. *Environ. Res. Lett.*, **13**, 094003, doi:10.1088/1748-9326/aad985.

Parker, D. J., P. Willetts, C. Birch, A. G. Turner, J. H. Marsham, C. M. Taylor, S. Kolusu, and G. M. Martin, 2016: The interaction of moist convection and mid-level dry air in the advance of the onset of the Indian monsoon. *Q. J. R. Meteorol. Soc.*, **142**, 2256–2272, doi:10.1002/qj.2815.

Pellicciotti, F., B. Brock, U. Strasser, P. Burlando, M. Funk, and J. Corripio, 2005: An enhanced temperature-index glacier melt model including the shortwave radiation balance: development and testing for Haut Glacier d'Arolla, Switzerland. *J. Glaciol.*, **51**, 573–587, doi:10.3189/172756505781829124.

——, T. Raschle, T. Huerlimann, M. Carenzo, and P. Burlando, 2011: Transmission of solar

radiation through clouds on melting glaciers: a comparison of parameterizations and their impact on melt modelling. *J. Glaciol.*, **57**, 367–381, doi:https://doi.org/10.3189/002214311796406013.

——, C. Buergi, W. W. Immerzeel, M. Konz, and A. B. Shrestha, 2012: Challenges and Uncertainties in Hydrological Modeling of Remote Hindu Kush-Karakoram-Himalayan (HKH) Basins: Suggestions for Calibration Strategies. *Mt. Res. Dev.*, **32**, 39–50, doi:10.1659/MRD-JOURNAL-D-11-00092.1.

——, C. Stephan, E. Miles, S. Herreid, W. W. Immerzeel, and T. Bolch, 2015: Mass-balance changes of the debris-covered glaciers in the Langtang Himal, Nepal, from 1974 to 1999. *J. Glaciol.*, **61**, 373–386, doi:10.3189/2015JoG13J237.

Pepin, N., D. Benham, and K. Taylor, 1999: Modeling Lapse Rates in the Maritime Uplands of Northern England: Implications for Climate Change. *Arctic, Antarct. Alp. Res.*, **31**, 151–164, doi:10.1080/15230430.1999.12003292.

——, and Coauthors, 2015: Elevation-dependent warming in mountain regions of the world. *Nat. Clim. Chang.*, **5**, 424–430, doi:10.1038/nclimate2563.

Pepin, N. C., E. E. Maeda, and R. Williams, 2016: Use of remotely sensed land surface temperature as a proxy for air temperatures at high elevations: Findings from a 5000m elevational transect across Kilimanjaro. *J. Geophys. Res. Atmos.*, **121**, 9998–10015, doi:10.1002/2016JD025497.

Peters-Lidard, C. D., and Coauthors, 2017: Scaling, similarity, and the fourth paradigm for hydrology. *Hydrol. Earth Syst. Sci.*, **21**, 3701–3713, doi:10.5194/hess-2016-695.

Platnick, S., M. D. King, S. A. Ackerman, W. P. Menzel, B. A. Baum, J. C. Riédi, and R. A. Frey, 2003: The MODIS Cloud Products: Algorithms and Examples from Terra. *IEEE Trans. Geosci. Remote Sens.*, **41**, doi:10.1109/TGRS.2002.808301.

Plummer, M. A., and F. M. Phillips, 2003: A 2-D numerical model of snow/ice energy balance and ice flow for paleoclimatic interpretation of glacial geomorphic features. *Quat. Sci. Rev.*, **22**, 1389–1406, doi:10.1016/S0277-3791(03)00081-7.

Prasch, M., W. Mauser, and M. Weber, 2013: Quantifying present and future glacier melt-water contribution to runoff in a central Himalayan river basin. *Cryosph.*, **7**, 889–904, doi:10.5194/tc-7-889-2013.

Prein, A. F., G. J. Holland, R. M. Rasmussen, J. Done, K. Ikeda, M. P. Clark, and C. H. Liu, 2012: Importance of Regional Climate Model Grid Spacing for the Simulation of Heavy Precipitation in the Colorado Headwaters. *J. Clim.*, **26**, 4848–4857, doi:10.1175/JCLI-D-12-00727.1.

Qu, X., and A. Hall, 2014: On the persistent spread in snow-albedo feedback. *Clim. Dyn.*, **42**, 69–81, doi:10.1007/s00382-013-1774-0.

Quincey, D. J., L. Copland, C. Mayer, M. Bishop, A. Luckman, and M. Belo, 2009: Ice

velocity and climate variations for Baltoro Glacier, Pakistan. *J. Glaciol.*, **55**, 1061–1071, doi:10.3189/002214309790794913.

——, M. Braun, N. F. Glasser, M. P. Bishop, K. Hewitt, and A. Luckman, 2011: Karakoram glacier surge dynamics. *Geophys. Res. Lett.*, **38**, L18504, 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**, 1288–1300, doi:10.1002/2015JF003515.

Ragettli, S., and F. Pellicciotti, 2012: Calibration of a physically based, spatially distributed hydrological model in a glacierized basin: On the use of knowledge from glaciometeorological processes to constrain model parameters. *Water Resour. Res.*, **48**, W03509, doi:10.1029/2011WR010559.

——, ——, R. Bordoy, and W. W. Immerzeel, 2013: Sources of uncertainty in modeling the glaciohydrological response of a Karakoram watershed to climate change. *Water Resour. Res.*, **49**, 6048–6066, doi:10.1002/wrcr.20450.

——, G. Cortés, J. Mcphee, and F. Pellicciotti, 2014: An evaluation of approaches for modelling hydrological processes in high-elevation, glacierized Andean watersheds. *Hydrol. Process.*, **28**, 5674–5695, doi:10.1002/hyp.10055.

——, and Coauthors, 2015: Unraveling the hydrology of a Himalayan catchment through integration of high resolution in situ data and remote sensing with an advanced simulation model. *Adv. Water Resour.*, **78**, 94–111, doi:10.1016/j.advwatres.2015.01.013. http://dx.doi.org/10.1016/j.advwatres.2015.01.013.

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.*, **113**, 9222–9227, doi:10.1073/pnas.1606526113.

Raleigh, M. S., and J. D. Lundquist, 2012: Comparing and combining SWE estimates from the SNOW-17 model using PRISM and SWE reconstruction. *Water Resour. Res.*, **48**, W01506, doi:10.1029/2011WR010542.

——, and M. P. Clark, 2014: Are temperature-index models appropriate for assessing climate change impacts on snowmelt? *82nd Annual Western Snow Conference*, Durango, Colorado, Proceedings of the Western Snow Conference, 45–51 https://westernsnowconference.org/sites/westernsnowconference.org/PDFs/2014Raleigh.pdf.

Ramanathan, V., P. J. Crutzen, J. T. Kiehl, and D. Rosenfeld, 2001: Aerosols, Climate and the Hydrological Cycle. *Science (80-.).*, **294**, 2119–2125, doi:10.1126/science.1064034.

Ramanathan, V., M. V Ramana, G. Roberts, D. Kim, C. Corrigan, C. Chung, and D. Winker, 2007: Warming trends in Asia amplified by brown cloud solar absorption. *Nature*, **448**, 575–579, doi:10.1038/nature06019.

Rangwala, I., E. Sinsky, and J. R. Miller, 2013: Amplified warming projections for high altitude regions of the northern hemisphere mid-latitudes from CMIP5 models. *Environ. Res. Lett.*, **8**, 024040, doi:10.1088/1748-9326/8/2/024040.

——, ——, and ——, 2016: Variability in projected elevation dependent warming in boreal midlatitude winter in CMIP5 climate models and its potential drivers. *Clim. Dyn.*, **46**, 2115–2122, doi:10.1007/s00382-015-2692-0. "http://dx.doi.org/10.1007/s00382-015-2692-0.

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

Rasmussen, R., and Coauthors, 2011: High-Resolution Coupled Climate Runoff Simulations of Seasonal Snowfall over Colorado: A Process Study of Current and Warmer Climate. *J. Clim.*, **24**, 3015–3048, doi:10.1175/2010JCLI3985.1.

Rasouli, K., J. W. Pomeroy, J. R. Janowicz, S. K. Carey, and T. J. Williams, 2014: Hydrological sensitivity of a northern mountain basin to climate change. *Hydrol. Process.*, **28**, 4191–4208, doi:10.1002/hyp.10244.

Rees, H. G., and D. N. Collins, 2006: Regional differences in response of flow in glacier-fed Himalayan rivers to climatic warming. **2169**, 2157–2169, doi:10.1002/hyp.

Reggiani, P., and T. H. M. Rientjes, 2015: A reflection on the long-term water balance of the Upper Indus Basin. *Hydrol. Res.*, **46**, 446–462, doi:10.2166/nh.2014.060.

Reggiani, P., G. Coccia, and B. Mukhopadhyay, 2016: Predictive Uncertainty Estimation on a Precipitation and Temperature Reanalysis Ensemble for Shigar Basin, Central Karakoram. *Water*, **8**, 1–25, doi:10.3390/w8060263.

——, B. Mukhopadhyay, T. H. M. Rientjes, and A. Khan, 2017: A joint analysis of river runoff and meteorological forcing in the Karakoram, upper Indus Basin. *Hydrol. Process.*, **31**, 409–430, doi:10.1002/hyp.11038.

Reid, T. D., and B. W. Brock, 2010: An energy-balance model for debris-covered glaciers including heat conduction through the debris layer. *J. Glaciol.*, **56**, 903–916, doi:10.3189/002214310794457218.

Reid, T. D., and B. W. Brock, 2014: Assessing ice-cliff backwasting and its contribution to total ablation of debris-covered Miage glacier, Mont Blanc massif, Italy. *J. Glaciol.*, **60**, 3–13, doi:10.3189/2014JoG13J045.

Réveillet, M., C. Vincent, D. Six, and A. Rabatel, 2017: Which empirical model is best suited to simulate glacier mass balances? *J. Glaciol.*, **63**, 39–54, doi:10.1017/jog.2016.110.

——, and Coauthors, 2018: Relative performance of empirical and physical models in assessing the seasonal and annual glacier surface mass balance of Saint-Sorlin Glacier (French Alps). *Cryosph.*, **12**, 1367–1386, doi:10.5194/tc-12-1367-2018.

Rice, R., R. C. Bales, T. H. Painter, and J. Dozier, 2011: Snow water equivalent along elevation gradients in the Merced and Tuolumne River basins of the Sierra Nevada. *Water Resour. Res.*, **47**, W08515, doi:10.1029/2010WR009278.

Ridley, J., A. Wiltshire, and C. Mathison, 2013: More frequent occurrence of westerly disturbances in Karakoram up to 2100. *Sci. Total Environ.*, **468-469**, S31–S35, doi:10.1016/j.scitotenv.2013.03.074. http://dx.doi.org/10.1016/j.scitotenv.2013.03.074.

Rienecker, M. M., and Coauthors, 2011: MERRA: NASA' s Modern-Era Retrospective Analysis for Research and Applications. *J. Clim.*, **24**, 3624–3648, doi:10.1175/JCLI-D-11-00015.1.

Rohrer, M., N. Salzmann, M. Stoffel, and A. V Kulkarni, 2013: Missing (in-situ) snow cover data hampers climate change and runoff studies in the Greater Himalayas. *Sci. Total Environ.*, **468-469**, S60–S70, doi:10.1016/j.scitotenv.2013.09.056. http://dx.doi.org/10.1016/j.scitotenv.2013.09.056.

Ruiz-Arias, J. A., C. Arbizu-Barrena, F. J. Santos-Alamillos, J. Tovar-Pescador, and D. Pozo-Vázquez, 2016: Assessing the Surface Solar Radiation Budget in the WRF Model: A Spatiotemporal Analysis of the Bias and Its Causes. *Mon. Weather Rev.*, **144**, 703–711, doi:10.1175/MWR-D-15-0262.1.

Rutter, N., and Coauthors, 2009: Evaluation of forest snow processes models (SnowMIP2). *J. Geophys. Res.*, **114**, D06111, doi:10.1029/2008JD011063.

Saeed, S., W. Müller, S. Hagemann, and D. Jacob, 2011: Circumglobal wave train and the summer monsoon over northwestern India and Pakistan: the explicit role of the surface heat low. *Clim. Dyn.*, **37**, 1045–1060, doi:10.1007/s00382-010-0888-x.

Saha, A., S. Ghosh, A. S. Sahana, and E. P. Rao, 2014: Failure of CMIP5 climate models in simulating post-1950 decreasing trend of Indian monsoon. *Geophys. Res. Lett.*, **41**, 7323–7330, doi:10.1002/2014GL061573.

Saha, S., and Coauthors, 2010: The NCEP Climate Forecast System Reanalysis. *Bull. Am. Meteorol. Soc.*, **91**, 1015–1058, doi:10.1175/2010BAMS3001.1.

Saha, S. K., K. Sujith, S. Pokhrel, H. S. Chaudhari, and A. Hazra, 2017: Effects of multilayer snow scheme on the simulation of snow: Offline Noah and coupled with NCEP CFSv2. *J. Adv. Model. Earth Syst.*, **9**, 271–290, doi:10.1002/2016MS000845.

Salomonson, V. V, and I. Appel, 2004: Estimating fractional snow cover from MODIS using the normalized difference snow index. *Remote Sens. Environ.*, **89**, 351–360, doi:10.1016/j.rse.2003.10.016.

Saloranta, T. M., 2012: Simulating snow maps for Norway: description and statistical evaluation of the seNorge snow model. *Cryosph.*, **6**, 1323–1337, doi:10.5194/tc-6-1323-2012.

Salvucci, G. D., and D. Entekhabi, 1994: Comparison of the Eagleson balance model with

numerical simulations. *Water Resour. Res.*, **30**, 2751–2757, doi:10.1029/94WR00949.

von Salzen, K., and Coauthors, 2013: The Canadian Fourth Generation Atmospheric Global Climate Model (CanAM4). Part I: Representation of Physical Processes. *Atmosphere-Ocean*, **51**, 104–125, doi:10.1080/07055900.2012.755610.

Sattar, E., J. Robison, and D. McCool, 2017: Evolution of Water Institutions in the Indus River Basin: Reflections from the Law of the Colorado River. *Univ. Michigan J. Law Reform*, 1–69, doi:10.2139/ssrn.3023589.

Savenije, H. H. G., and M. Hrachowitz, 2017: HESS Opinions "Catchments as meta-organisms – a new blueprint for hydrological modelling." *Hydrol. Earth Syst. Sci.*, **21**, 1107–1116, doi:10.5194/hess-21-1107-2017.

de Scally, F. A., 1992: Influence of avalanche snow transport on snowmelt runoff. *J. Hydrol.*, **137**, 73–97, doi:10.1016/0022-1694(92)90049-2.

De Scally, F. A., 1994: Relative importance of snow accumulation and monsoon rainfall data for estimating annual runoff, Jhelum basin, Pakistan. *Hydrol. Sci. J.*, **39**, 199–216, doi:10.1080/02626669409492738.

de Scally, F. A., 1997: Deriving Lapse Rates of Slope Air Temperature for Meltwater Runoff Modeling in Subtropical Mountains: An Example from the Punjab Himalaya, Pakistan. *Mt. Res. Dev.*, **17**, 353–362, doi:10.2307/3674024.

de Scally, F. A., and I. S. Gardner, 1989: Evaluation of Avalanche-Mass Determination Approaches: An Example from the Himalaya, Pakistan. *J. Glaciol.*, **35**, 248–252, doi:10.3189/S0022143000004573.

de Scally, F. A., and J. S. Gardner, 1990: Ablation of Avalanched and Undisturbed Snow, Himalaya Mountains, Pakistan. *Water Resour. Res.*, **26**, 2757–2767, doi:10.1029/WR026i011p02757.

Scherler, D., and M. R. Strecker, 2012: Large surface velocity fluctuations of Biafo Glacier, central Karakoram, at high spatial and temporal resolution from optical satellite images. *J. Glaciol.*, **58**, 569–580, doi:10.3189/2012JoG11J096.

——, B. Bookhagen, and M. R. Strecker, 2011: Spatially variable response of Himalayan glaciers to climate change affected by debris cover. *Nat. Geosci.*, **4**, 156–159, doi:10.1038/ngeo1068. http://dx.doi.org/10.1038/ngeo1068.

Schiemann, R., D. Lüthi, and C. Schär, 2009: Seasonality and Interannual Variability of the Westerly Jet in the Tibetan Plateau Region. *J. Clim.*, **22**, 2940–2957, doi:10.1175/2008JCLI2625.1.

Schleussner, C.-F., and Coauthors, 2016: Differential climate impacts for policy-relevant limits to global warming: the case of 1.5°C and 2°C. *Earth Syst. Dyn.*, **7**, 327–351, doi:10.5194/esd-7-327-2016.

Serinaldi, F., C. G. Kilsby, and F. Lombardo, 2018: Untenable nonstationarity: An assessment of the fitness for purpose of trend tests in hydrology. *Adv. Water Resour.*, **111**, 132–155, doi:10.1016/j.advwatres.2017.10.015. https://doi.org/10.1016/j.advwatres.2017.10.015.

Sharif, M., D. R. Archer, H. J. Fowler, and N. Forsythe, 2013: Trends in timing and magnitude of flow in the Upper Indus Basin. *Hydrol. Earth Syst. Sci.*, **17**, 1503–1516, doi:10.5194/hess-17-1503-2013.

Shrestha, M., T. Koike, Y. Hirabayashi, Y. Xue, L. Wang, G. Rasul, and B. Ahmad, 2015: Integrated simulation of snow and glacier melt in water and energy balance-based, distributed hydrological modeling framework at Hunza River Basin of Pakistan Karakoram region. *J. Geophys. Res. Atmos.*, **120**, 4889–4919, doi:10.1002/2014JD022666.

Sicart, J. E., J. W. Pomeroy, R. L. H. Essery, and D. Bewley, 2006: Incoming longwave radiation to melting snow: observations, sensitivity and estimation in northern environments. *Hydrol. Process.*, **20**, 3697–3708, doi:10.1002/hyp.6383.

Silverman, N. L., M. P. Maneta, S.-H. Chen, and J. T. Harper, 2013: Dynamically downscaled winter precipitation over complex terrain of the Central Rockies of Western Montana, USA. *Water Resour. Res.*, **49**, 458–470, doi:10.1029/2012WR012874.

Sinclair, S., and G. G. S. Pegram, 2010: A comparison of ASCAT and modelled soil moisture over South Africa, using TOPKAPI in land surface mode. *Hydrol. Earth Syst. Sci.*, **14**, 613–626, doi:10.5194/hess-14-613-2010.

——, and ——, 2013: A sensitivity assessment of the TOPKAPI model with an added infiltration module. *J. Hydrol.*, **479**, 100–112, doi:10.1016/j.jhydrol.2012.11.061. http://dx.doi.org/10.1016/j.jhydrol.2012.11.061.

Singh, P., and N. Kumar, 1996: Determination of snowmelt factor in the Himalayan region. *Hydrol. Sci. J.*, **41**, 301–310, doi:10.1080/02626669609491504. http://www.tandfonline.com/doi/abs/10.1080/02626669609491504.

Singh, P., N. Kumar, and M. Arora, 2000: Degree-day factors for snow and ice for Dokriani Glacier, Garhwal Himalayas. *J. Hydrol.*, **235**, 1–11, doi:10.1016/S0022-1694(00)00249-3.

Skamarock, W. C., and Coauthors, 2008: *A Description of the Advanced Research WRF Version 3*. NCAR Technical Note NCAR/TN–475+STR, 1-113 pp.

Smiraglia, C., G. Diolaiuti, C. Mayer, A. Lambrecht, and M. Belo, 2006: Glaciological characteristics of the ablation zone of Baltoro glacier , Karakoram , Pakistan. 123–131.

Smith, T., and B. Bookhagen, 2018: Changes in seasonal snow water equivalent distribution in High Mountain Asia (1987 to 2009). *Sci. Adv.*, **4**, e1701550, doi:10.1126/sciadv.1701550.

Soncini, A., and Coauthors, 2014: Future Hydrological Regimes in the Upper Indus Basin: A

Case Study from a High-Altitude Glacierized Catchment. *J. Hydrometeorol.*, **16**, 306–326, doi:10.1175/JHM-D-14-0043.1.

Sperber, K. R., H. Annamalai, I.-S. Kang, A. Kitoh, A. Moise, A. Turner, B. Wang, and T. Zhou, 2013: The Asian summer monsoon: an intercomparison of CMIP5 vs. CMIP3 simulations of the late 20th century. *Clim. Dyn.*, **41**, 2711–2744, doi:10.1007/s00382-012-1607-6.

Stevens, B., and Coauthors, 2013: Atmospheric component of the MPI-M Earth System Model: ECHAM6. *J. Adv. Model. Earth Syst.*, **5**, 146–172, doi:10.1002/jame.20015.

Stigter, E. E., M. Litt, J. F. Steiner, P. N. J. Bonekamp, J. M. Shea, M. F. P. Bierkens, and W. W. Immerzeel, 2018: The Importance of Snow Sublimation on a Himalayan Glacier. *Front. Earth Sci.*, **6**, 108, doi:10.3389/feart.2018.00108.

Strasser, U., M. Bernhardt, M. Weber, G. E. Liston, and W. Mauser, 2008: Is snow sublimation important in the alpine water balance? *Cryosph.*, **2**, 53–66, doi:10.5194/tc-2-53-2008.

Su, F., X. Duan, D. Chen, Z. Hao, and L. Cuo, 2013: Evaluation of the Global Climate Models in the CMIP5 over the Tibetan Plateau. *J. Clim.*, **26**, 3187–3208, doi:10.1175/JCLI-D-12-00321.1.

Syed, F. S., F. Giorgi, J. S. Pal, and M. P. King, 2006: Effect of remote forcings on the winter precipitation of central southwest Asia part 1: observations. *Theor. Appl. Climatol.*, **86**, 147–160, doi:10.1007/s00704-005-0217-1.

——, ——, ——, and K. Keay, 2010: Regional climate model simulation of winter climate over Central–Southwest Asia, with emphasis on NAO and ENSO. *Int. J. Climatol.*, **30**, 220–235, doi:10.1002/joc.1887.

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

Tahir, A. A., P. Chevallier, Y. Arnaud, L. Neppel, and B. Ahmad, 2011b: Modeling snowmelt-runoff under climate scenarios in the Hunza River basin, Karakoram Range, Northern Pakistan. *J. Hydrol.*, **409**, 104–117, doi:10.1016/j.jhydrol.2011.08.035. http://dx.doi.org/10.1016/j.jhydrol.2011.08.035.

——, ——, ——, M. Ashraf, and M. T. Bhatti, 2015: Snow cover trend and hydrological characteristics of the Astore River basin (Western Himalayas) and its comparison to the Hunza basin (Karakoram region). *Sci. Total Environ.*, **505**, 748–761, doi:10.1016/j.scitotenv.2014.10.065. http://dx.doi.org/10.1016/j.scitotenv.2014.10.065.

Tarasova, L., M. Knoche, J. Dietrich, and R. Merz, 2016: Effects of input discretization, model complexity, and calibration strategy on model performance in a data-scarce glacierized catchment in Central Asia. *Water Resour. Res.*, **52**, 4674–4699, doi:10.1002/2015WR018551.

Tarboton, D. G., T. G. Chowdhury, and T. H. Jackson, 1995: A spatially distributed energy balance snowmelt model. *Biogeochemistry of Seasonally Snow-Covered Catchments (Proceedings of a Boulder Symposium)*, Boulder, IAHS, 141–155.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An overview of CMIP5 and the experiment design. *Bull. Am. Meteorol. Soc.*, **93**, 485–498, doi:10.1175/BAMS-D-11-00094.1.

Tebaldi, C., and R. Knutti, 2018: Evaluating the accuracy of climate change pattern emulation for low warming targets. *Environ. Res. Lett.*, **13**, 055006, doi:10.1088/1748-9326/aabef2.

Thompson, G., M. Tewari, K. Ikeda, S. Tessendorf, C. Weeks, J. Otkin, and F. Kong, 2016: Explicitly-coupled cloud physics and radiation parameterizations and subsequent evaluation in WRF high-resolution convective forecasts. *Atmos. Res.*, **168**, 92–104, doi:10.1016/j.atmosres.2015.09.005.

Tiwari, S., S. C. Kar, and R. Bhatla, 2017: Atmospheric moisture budget during winter seasons in the western Himalayan region. *Clim. Dyn.*, **48**, 1277–1295, doi:10.1007/s00382-016-3141-4. "http://dx.doi.org/10.1007/s00382-016-3141-4.

Todini, E., 1995: New trends in modelling soil processes from hillslope to GCM scales. *The role of water and the hydrological cycle in global change*, H. Oliver and S. Oliver, Eds., Springer, Berlin Heidelberg, 317–347.

Todini, E., 1996: The ARNO rainfall-runoff model. *J. Hydrol.*, **175**, 339–382, doi:10.1016/S0022-1694(96)80016-3.

Tomasi, E., L. Giovannini, D. Zardi, and M. de Franceschi, 2017: Optimization of Noah and Noah_MP WRF Land Surface Schemes in Snow-Melting Conditions over Complex Terrain. *Mon. Weather Rev.*, **145**, 4727–4745, doi:10.1175/MWR-D-16-0408.1.

UN, 2017: *World Population Prospects: The 2017 Revision, Key Findings and Advance Tables*. 1-53 pp.

UNDP, 2006: *Human Development Report 2006. Beyond scarcity: Power, poverty and the global water crisis*. 1-422 pp.

——, 2016: Water Security in Pakistan: Issues and Challenges. *Dev. Advocate Pakistan*, **3**, 1–39.

Uppala, S. M., and Coauthors, 2005: The ERA-40 re-analysis. *Q. J. R. Meteorol. Soc.*, **131**, 2961–3012, doi:10.1256/qj.04.176.

Vincent, C., and D. Six, 2013: Relative contribution of solar radiation and temperature in enhanced temperature-index melt models from a case study at Glacier de Saint-Sorlin, France. *Ann. Glaciol.*, **54**, 11–17, doi:10.3189/2013AoG63A301.

Vincent, C., and Coauthors, 2016: Reduced melt on debris-covered glaciers: investigations

from Changri Nup Glacier , Nepal. *Cryosph.*, **10**, 1845–1858, doi:10.5194/tc-10-1845-2016.

Viviroli, D., and Coauthors, 2011: Climate change and mountain water resources: overview and recommendations for research, management and policy. *Hydrol. Earth Syst. Sci.*, **15**, 471–504, doi:10.5194/hess-15-471-2011.

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, doi:10.3189/S0260305500008053.

Wan, Z., Y. Zhang, Q. Zhang, and Z.-L. Li, 2004: Quality assessment and validation of the MODIS global land surface temperature. *Int. J. Remote Sens.*, **25**, 261–274, doi:10.1080/0143116031000116417.

Wang, Y., S. Hou, B. Huai, W. An, H. Pang, and Y. Liu, 2018: Glacier anomaly over the western Kunlun Mountains, Northwestern Tibetan Plateau, since the 1970s. *J. Glaciol.*, **64**, 624–636, doi:10.1017/jog.2018.53.

Wang, Z., X. Zeng, and M. Decker, 2010: Improving snow processes in the Noah land model. *J. Geophys. Res.*, **115**, D20108, doi:10.1029/2009JD013761.

Wang, Z., and Coauthors, 2012: Evaluation of Moderate-resolution Imaging Spectroradiometer (MODIS) snow albedo product (MCD43A) over tundra. *Remote Sens. Environ.*, **117**, 264–280, doi:10.1016/j.rse.2011.10.002.

Wang, Z., G. Li, and S. Yang, 2017: Origin of Indian summer monsoon rainfall biases in CMIP5 multimodel ensemble. *Clim. Dyn.*, **51**, 755–768, doi:10.1007/s00382-017-3953-x. http://dx.doi.org/10.1007/s00382-017-3953-x.

Waqas, A., and H. Athar, 2018: Observed diurnal temperature range variations and its association with observed cloud cover in northern Pakistan. *Int. J. Climatol.*, **38**, 3323–3336, doi:10.1002/joc.5503.

Watanabe, M., and Coauthors, 2010: Improved Climate Simulation by MIROC5: Mean States, Variability, and Climate Sensitivity. *J. Clim.*, **23**, 6312–6335, doi:10.1175/2010JCLI3679.1.

Wehner, M. F., and Coauthors, 2014: The effect of horizontal resolution on simulation quality in the Community Atmospheric Model, CAM5.1. *J. Adv. Model. Earth Syst.*, **6**, 980–997, doi:10.1002/2013MS000276.

Wei, W., R. Zhang, M. Wen, and S. Yang, 2017: Relationship between the Asian Westerly Jet Stream and Summer Rainfall over Central Asia and North China: Roles of the Indian Monsoon and the South Asian High. *J. Clim.*, **30**, 537–552, doi:10.1175/JCLI-D-15-0814.1.

Wen, J., Q. Liu, Q. Xiao, Q. Liu, D. You, D. Hao, S. Wu, and X. Lin, 2018: Characterizing Land Surface Anisotropic Reflectance over Rugged Terrain: A Review of Concepts and

Recent Developments. *Remote Sens.*, **10**, 1–30, doi:10.3390/rs10030370.

Wenbin, Z., L. Aifeng, and J. Shaofeng, 2013: Estimation of daily maximum and minimum air temperature using MODIS land surface temperature products. *Remote Sens. Environ.*, **130**, 62–73, doi:10.1016/j.rse.2012.10.034. http://dx.doi.org/10.1016/j.rse.2012.10.034.

Wescoat, J. L., G. White, and K. Mitchell, 1991: Managing the lndus River basin in light of climate change: Four conceptual approaches. *Glob. Environ. Chang.*, **1**, 381–395, doi:10.1016/0959-3780(91)90004-D.

Wilby, R. L., and S. Dessai, 2010: Robust adaptation to climate change. *Weather*, **65**, 180–185, doi:10.1002/wea.543.

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

Wood, E. F., and Coauthors, 2011: Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth' s terrestrial water. *Water Resour. Res.*, **47**, W05301, doi:10.1029/2010WR010090.

Xu, J., Y. Gao, D. Chen, and T. Ou, 2017: Evaluation of global climate models for downscaling applications centred over the Tibetan Plateau. *Int. J. Climatol.*, **37**, 657–671, doi:10.1002/joc.4731.

Yadav, R. K., K. R. Kumar, and M. Rajeevan, 2009: Increasing influence of ENSO and decreasing influence of AO/NAO in the recent decades over northwest India winter precipitation. *J. Geophys. Res.*, **114**, D12112, doi:10.1029/2008JD011318.

——, J. H. Yoo, F. Kucharski, and M. A. Abid, 2010: Why Is ENSO Influencing Northwest India Winter Precipitation in Recent Decades? *J. Clim.*, **23**, 1979–1993, doi:10.1175/2009JCLI3202.1.

Yatagai, A., K. Kamiguchi, O. Arakawa, A. Hamada, N. Yasutomi, and A. Kitoh, 2012: APHRODITE: Constructing a Long-Term Daily Gridded Precipitation Dataset for Asia Based on a Dense Network of Rain Gauges. *Bull. Am. Meteorol. Soc.*, **93**, 1401–1415, doi:10.1175/BAMS-D-11-00122.1.

Young, G. J., and K. Hewitt, 1990: Hydrology research in the upper Indus basin, Karakoram Himalaya, Pakistan. *Hydrology of Mountainous Areas (Proceedings of the Štrbské Pleso Workshop, Czechoslovakia, June 1989)*, 139–152.

Zawahri, N., and D. Michel, 2018: Assessing the Indus Waters Treaty from a comparative perspective. *Water Int.*, 1–17, doi:10.1080/02508060.2018.1498994. https://doi.org/10.1080/02508060.2018.1498994.

Zhang, H., F. Zhang, G. Zhang, T. Che, and W. Yan, 2018: How Accurately Can the Air Temperature Lapse Rate Over the Tibetan Plateau Be Estimated From MODIS LSTs? *J. Geophys. Res. Atmos.*, **123**, 3943–3960, doi:10.1002/2017JD028243.

Zhou, Y., Z. Li, and J. I. A. Li, 2017: Slight glacier mass loss in the Karakoram region during the 1970s to 2000 revealed by KH-9 images and SRTM DEM. *J. Glaciol.*, **63**, 331–342, doi:10.1017/jog.2016.142.

Appendix A

Supplementary Information for Chapter 3

A.1 Data and Methods

A.1.1 High Asia Refined Analysis (HAR)

In this section we provide additional background on the High Asia Refined Analysis (HAR) re-initialisation strategy (Maussion et al. 2014). Re-initialisation here refers to the approach in which multiple model runs covering consecutive time periods are performed, with each run having its own initialisation of atmospheric and land surface states. A long, continuous series of model results is thus generated by concatenating multiple runs that cover subsets of the overall period. This can be contrasted with continuous, long-term integrations based on a single initialisation of the atmospheric and land surface variables.

From preliminary testing (Maussion et al. 2011), Maussion et al. (2011, 2014) found that re-initialising the model on a daily basis outperformed weekly re-initialisation. They also concluded that spectral nudging would likely lead to unacceptable levels of drift in land surface and near-surface variables. This would have been problematic, as they aimed to simulate near-surface / boundary layer conditions as accurately as possible, in line with their intention for the HAR to be used to support offline hydrological and cryospheric modelling. The daily re-initialisation strategy they adopted is outlined in Section 3.3.1, but basically it involves performing a separate, independent WRF simulation for each day in the October 2000 to October 2014 period. For any given day, the atmospheric and land surface states required as initial conditions in the simulation are interpolated to the model grid from the driving FNL dataset. Atmospheric variables required include quantities like temperature and humidity, whereas land surface states needed include variables like snow water equivalent (SWE) and soil moisture.

The daily re-initialisation strategy substantially reduces the potential for excessive drift from initial states of both atmospheric and land surface variables derived from the FNL dataset. As described in Section 3.3.1, each simulation was for 36 hours, with the first 12 hours (i.e. the second half of the previous day) regarded as spin-up. This is common to all components of

the WRF model, i.e. atmosphere and land surface. This could be a limitation for slower responding land surface variables like soil moisture, but details on this are not reported by Maussion et al. (2014). Moreover, while simulating all the days independently enabled parallelisation and so computational tractability, it also means that the value added by the high resolution of the HAR does not feed into improved initial conditions each day (relative to the coarse FNL dataset).

These issues exemplify the compromises between competing objectives and feasibility in the HAR design. Despite this, the HAR re-initialisation and spin-up approach has already been demonstrated to be sufficient for a number of applications (Section 3.1), and the results in this study further support its utility and potential as an approach to characterising near-surface climate in the data-sparse Upper Indus Basin (UIB). Moreover, very similar re-initialisation and spin-up approaches have been adopted in other recent Himalayan modelling studies (e.g. Norris et al. 2015; Cannon et al. 2017). Nevertheless, future work could use sensitivity tests to further explore the implications of this approach and compare it with simulations allowing for longer spin-up periods, as discussed in Section 5.

A.1.2 Elevation Corrections

This section expands on the elevation corrections for the monthly mean temperature bias and incoming longwave radiation evaluations discussed in the main manuscript. The regressions used in correction are based on monthly climatologies of the 9 HAR grid cells surrounding each station, such that we get one gradient for each month of the year (i.e. 12 in total) for each station. Using monthly climatology as the basis for regression provides a balance between capturing intra-annual variation and ensuring stability in the calculations.

We summarise the results of the regressions of climatological monthly mean temperatures and elevation in Table A.1. The consistently high R^2 values demonstrate that elevation explains a very high proportion of variance in climatological monthly mean temperature in the 9 HAR grid cells surrounding each station. This is the case for all of the stations and throughout the year (monthly variation is shown as minima and maxima). The P values are all low and statistically significant at the 95% interval, with almost all significant at the 99% interval too. This suggests that the calculated gradients are robust for elevation corrections at this time aggregation.

Station	Gradient (°C/km)			R^2			P value		
	Avg	**Min**	**Max**	**Avg**	**Min**	**Max**	**Avg**	**Min**	**Max**
Askole	-6.92	-8.29	-5.71	0.96	0.87	1.00	0.000	0.000	0.000
Astore	-5.80	-6.92	-4.38	0.89	0.65	0.99	0.002	0.000	0.009
Concordia	-7.21	-8.69	-6.39	0.99	0.97	1.00	0.000	0.000	0.000
Gilgit	-6.58	-7.89	-5.67	0.99	0.95	1.00	0.000	0.000	0.000
Khunjerab	-6.75	-8.15	-5.64	0.85	0.70	1.00	0.001	0.000	0.005
Naltar	-6.88	-8.73	-6.09	0.98	0.93	1.00	0.000	0.000	0.000
Rama	-5.71	-7.19	-4.38	0.89	0.65	0.99	0.002	0.000	0.009
Rattu	-5.46	-6.92	-4.38	0.86	0.65	0.99	0.002	0.000	0.009
Skardu	-5.49	-7.37	-3.51	0.89	0.50	1.00	0.002	0.000	0.033
Urdukas	-6.78	-8.10	-5.62	0.98	0.95	1.00	0.000	0.000	0.000
Ushkore	-6.43	-8.09	-5.43	0.91	0.67	1.00	0.001	0.000	0.007
Yasin	-6.83	-8.75	-5.48	0.97	0.95	0.99	0.000	0.000	0.000
Ziarat	-7.16	-8.24	-5.84	0.92	0.78	0.99	0.000	0.000	0.002
Mean	**-6.46**	**-7.95**	**-5.27**	**0.93**	**0.79**	**1.00**	**0.00**	**0.00**	**0.01**

Table A.1 – Regression summary for climatological monthly mean temperatures and elevation. The average (avg), minimum (min) and maximum (max) aggregations summarise intra-annual (monthly) variation.

These gradients are only used in the evaluation of bias in the monthly mean temperature climatology in Section 3.4 (Figure 3.6a) in the main manuscript. To test the sensitivity of the bias calculations to our approach, we also calculated an elevation-adjusted climatology based on a constant lapse rate of -6.5°C/km. As shown in Figure A.1, this reduces the variation in bias compared with the monthly climatological lapse rate approach, but it does not ameliorate the cold bias or alter the seasonality of the bias. Therefore, we consider the bias pattern to be robust. If no elevation adjustment is made at all, the annual cycle of bias again remains similar but the overall magnitude of the bias increases, i.e. it becomes colder. For example,

while the HAR10 mean annual bias with an elevation correction using a constant -6.5°C/km is -3.9°C, this worsens to -8.5°C if elevation is not accounted for at all.

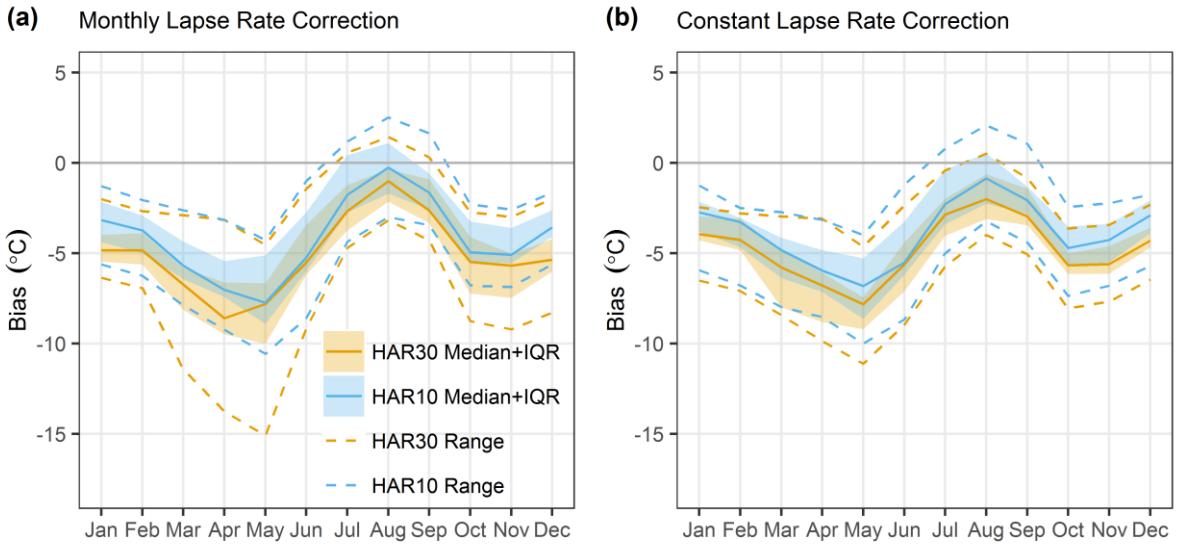


Figure A.1 – Comparison of monthly mean temperature biases at station locations using (a) monthly varying (climatological) lapse rate correction and (b) constant (-6.5°C/km) lapse rate correction.

In terms of longwave radiation, a relatively small performance gain is achieved with elevation correction using monthly climatological lapse rates. For example, mean annual bias in HAR10 decreases from 67 W/m^2 to 56 W/m^2, while HAR30 mean annual bias reduces from 50 W/m^2 to 41 W/m^2. Most of the improvement is evident in winter, spring and autumn, whereas in summer the strength of association with elevation decreases and so the performance gain is lesser, as shown in Figure A.2. This suggests that our conclusions in Section 4 on underestimation of incoming longwave radiation at this location, particularly in summer, hold both with and without the application of an elevation correction.

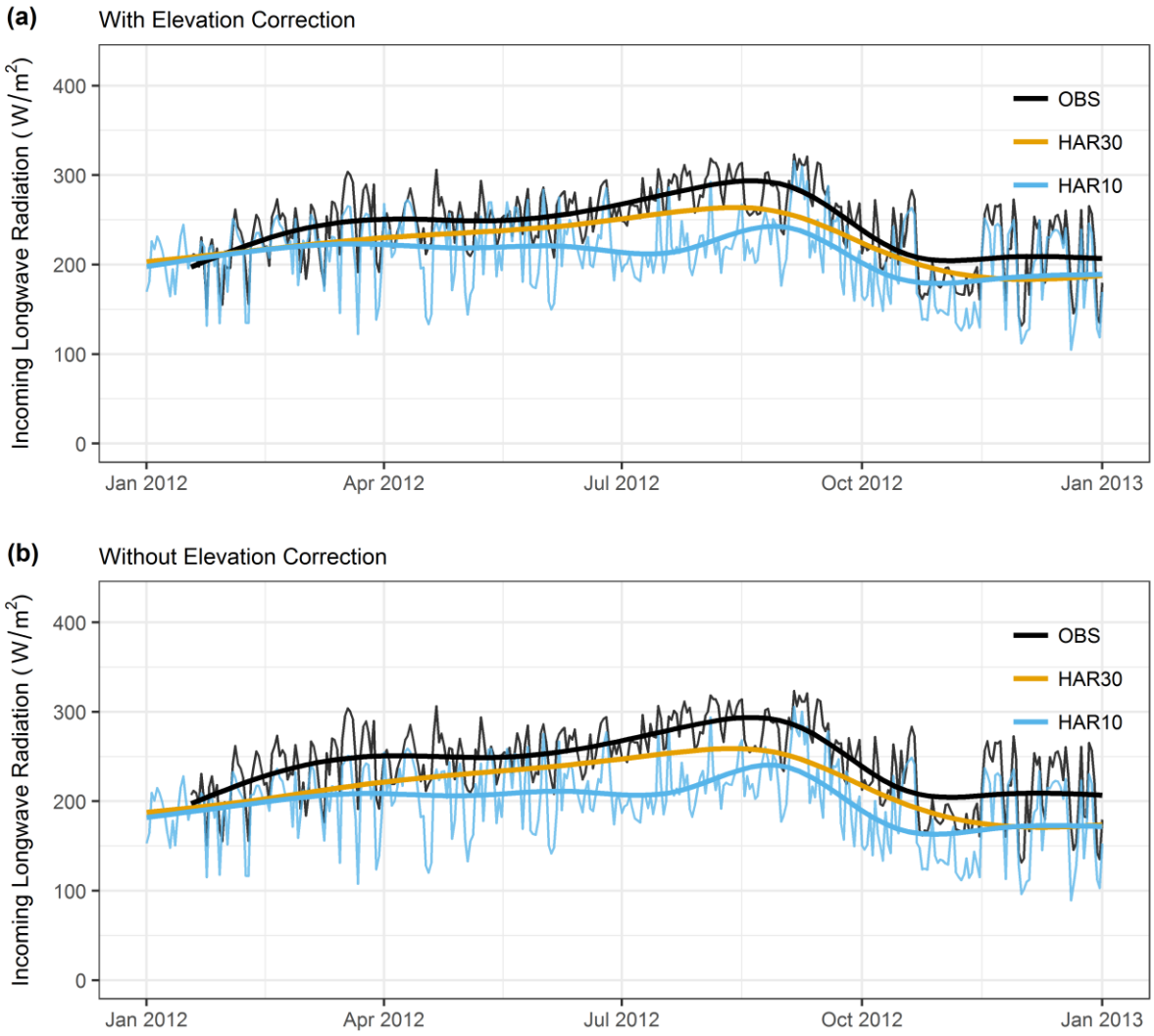


Figure A.2 – Comparison of incoming longwave radiation time series at Concordia AWS with (a) and without (a) elevation correction based on monthly climatological lapse rates.

A.1.3 Water Balance Calculations

This section provides some further discussion of why we compare mean annual effective precipitation to observed runoff, rather than HAR-simulated runoff. Essentially, the main reason follows from the daily re-initialisation strategy employed in the HAR. As described in Section 3.3.1 and Maussion et al. (2014), each day was initialised and run separately, which permitted parallelisation and therefore computational tractability of the project. One corollary of this approach is that SWE was initialised each day from the coarser FNL driving dataset. The accuracy of the SWE field is limited by the fact that insufficient data are available to produce a representative SWE analysis in this region, as well as the coarse resolution of the FNL data product. As so much of the runoff in the UIB originates from snow and ice melt,

simulated runoff from the Noah LSM becomes dominated by the limitations of the SWE field from the FNL initialisation dataset.

We determined this from a preliminary analysis of the HAR-computed runoff at a mean annual time scale, which showed HAR runoff to be very low relative to HAR effective precipitation (calculated as precipitation minus actual evapotranspiration). Consistent with this, we found the HAR peak SWE to be low relative to simulated snowfall. While there are multiple limitations to the representation of snow, glacier and hydrological processes in the Noah LSM, the daily initialisation of SWE from the FNL dataset appears dominant in shaping the underestimation of mean annual runoff.

In contrast, Figure 3b shows that mean annual HAR effective precipitation is more consistent with observed runoff, acknowledging the variation between sub-basins. This leads to the question of whether effective precipitation may be a more useful quantity to look at for establishing the plausibility of HAR precipitation for the water balance. Indeed, this is the main point that we wish to investigate here, as the primary purpose of our study is to assess the near-surface climate representation in the HAR. As discussed in Chapter 3, the utility of effective precipitation depends on the plausibility of the actual evapotranspiration (ET) estimates from the HAR. We find that HAR ET falls in the range of estimates from other data sources (see Table 3.4 and references in Section 3.4.1). This provides some empirical support for the HAR ET estimates, but we stress that the uncertainties in all estimates are large in this data-sparse region. Consistency between datasets is not a guide to their absolute accuracy, given that each dataset has its own errors and biases, but it still helps to provide some indication that HAR precipitation may be a useful starting point for further studies of the UIB and its water balance.

A.2 Precipitation

A.2.1 Seasonality Comparison

Figure A.3 compares HAR10 precipitation seasonality with selected global reanalyses and HAR10. Spatial means for the NWUIB area are used in this comparison, in line with the coarser resolution of the global reanalyses. This explains why the HAR10 series is slightly different to that plotted on Figure 3.2b, which is just based on station locations, although the important pattern of having a relatively high fraction of annual precipitation in the first few months of the year and a relatively low fraction in summer is the same. Figure A.3 shows that

ERA-Interim and NASA MERRA2 have quite similar precipitation climatologies, which are in reasonable agreement with the mean of the observations. In both of these reanalyses, more of the annual precipitation occurs in winter and spring, but there is a non-negligible contribution in the summer months, which is in line with the mean of the observations. The HAR10 line shows a starker annual cycle, as expected from Figure 3.2b, while JRA55 is almost an intermediate case, with a relatively high fraction in the early part of the year but still more summer precipitation compared with the HAR.

This indicates that there is reasonable consistency in the shape of the annual precipitation cycle in the global reanalyses examined here, which are in overall agreement with observations, although of course there is some degree of variation. It is interesting then that the HAR tends to under-predict the fraction of annual precipitation occurring in summer relative to both observations and global reanalyses.

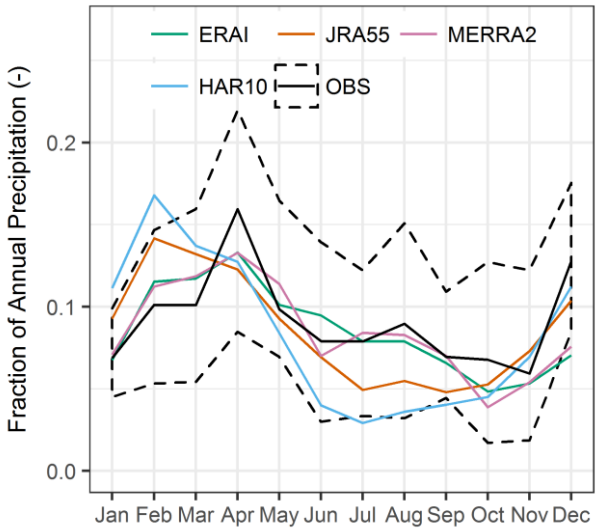


Figure A.3 – Comparison of precipitation seasonality between HAR10, selected global reanalyses and station observations for the NWUIB for the overlapping record period.

A.2.2 Inter-Annual Variability

Figure A.4a shows that, in addition to reproducing most mean annual observations (see Section 3.4.1), HAR10 matches observed inter-annual variability fairly well in most cases. The higher variation evident for HAR30 is a corollary of its larger mean precipitation. Normalising for this difference shows that both products have a similar coefficient of variation (CV) for each station (not shown). The HAR CV values are also of the correct order

of magnitude, but do not display particularly clear association with observed variation between stations. We extend the comparison in Figure A.4b by examining the (absolute) monthly variability, which is well captured in HAR10 and higher in HAR30, again owing to the latter's larger overall magnitude. Overall this supports Maussion et al.'s (2014) finding that the HAR contains a useful representation of inter-annual precipitation variability.

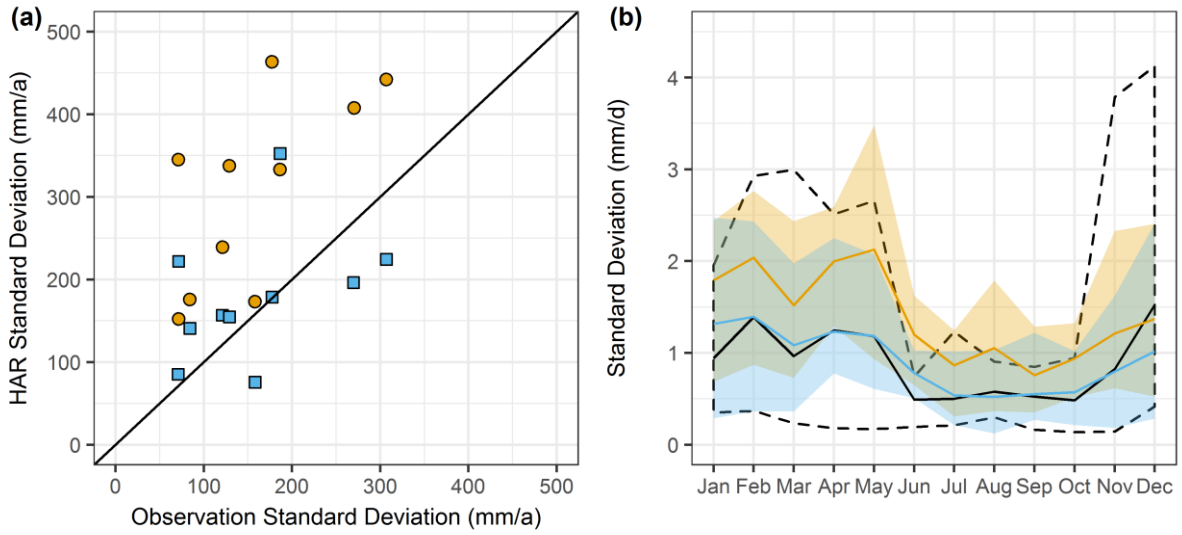


Figure A.4 – Comparison of HAR precipitation variability with station observations. Annual standard deviations are given in (a), while (b) shows monthly standard deviations. The solid line in (b) shows the mean for all stations, while shading shows the range.

A.2.3 Time Series

Monthly time series for the station locations are given in Figure A.5. These series clearly show the added value of the HAR10's high resolution in most cases. As expected from Figure 3.2, the greatest differences are at the Khunjerab, Ushkore and Ziarat stations.

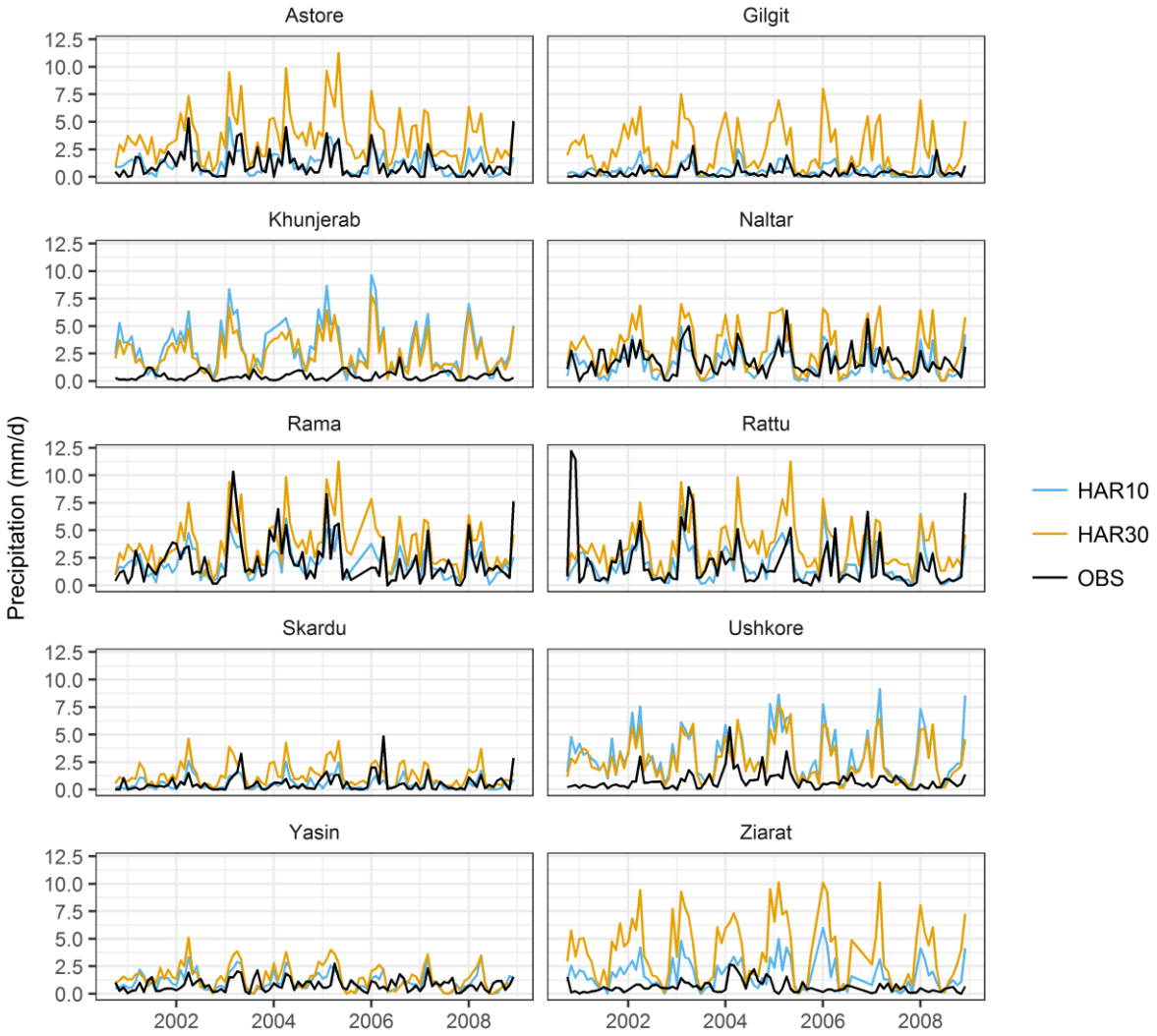


Figure A.5 – Comparison of monthly precipitation time series between the HAR and observations at station locations.

A.2.4 Mean Annual Precipitation Maps

To illustrate the effects of resolution on precipitation skill, we plot mean annual precipitation for the HAR10 and the HAR30 in Figure A.6. This is similar to maps shown in Maussion et al. (2014), but limited to the UIB extent. Figure A.6 shows that both the HAR10 and HAR30 generally agree on the major zones of orographic enhancement in the front ranges to the south-west of the UIB, as well as in the Hunza, Shigar and northern parts of the Shyok basin. However, the differences between Figure A.6a and Figure A.6b highlight the enhanced ability of the HAR10 to represent topographic variation and its influences on precipitation contrasts between valleys and ridges. This helps to explain why HAR30 provides an overestimate relative to observed precipitation and runoff (see Section 3.4.1).

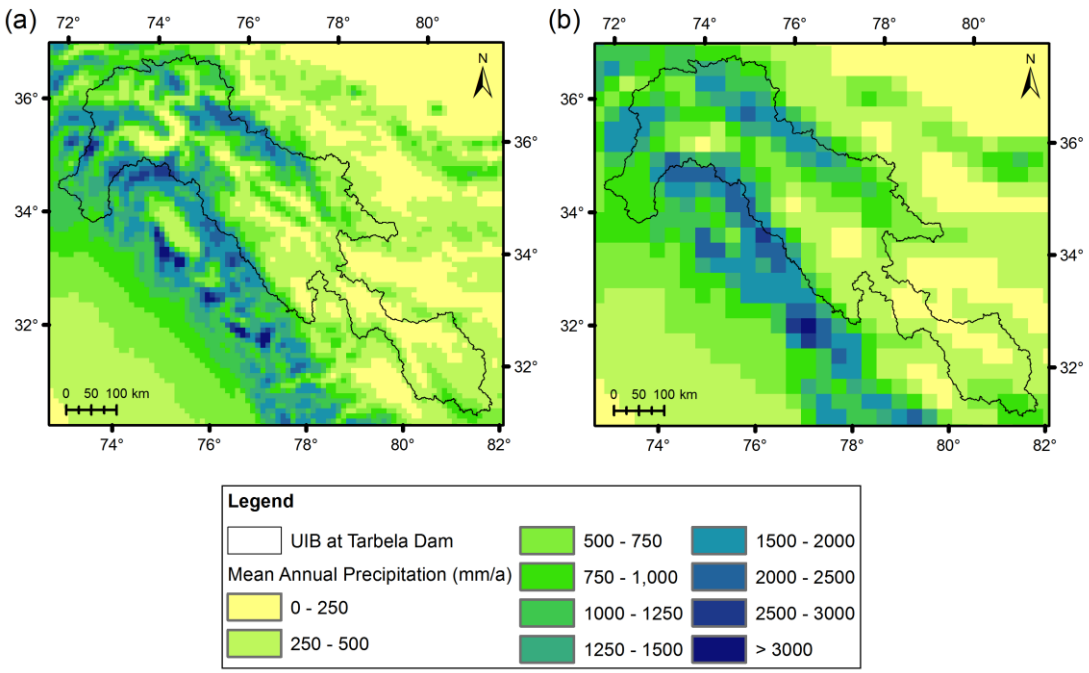


Figure A.6 – Maps of mean annual precipitation for (a) HAR10 and (b) HAR30.

A.2.5 Spatial and Vertical Gradients

To complement the evaluation of spatial and vertical precipitation gradients in the main text (Section 3.4.1), we analyse additional published vertical profiles here. In addition to the small number of observation-based vertical profiles available in other studies (Hewitt 2014; Winiger et al. 2005), these include model-based precipitation reconstructions from glacier mass balance. The profile attributed to Lutz et al. (2016) was derived by the present authors for the NWUIB domain based on the historical baseline precipitation datasets released with their paper, which are in turn based on Immerzeel et al.'s (2015) model-based reconstruction.

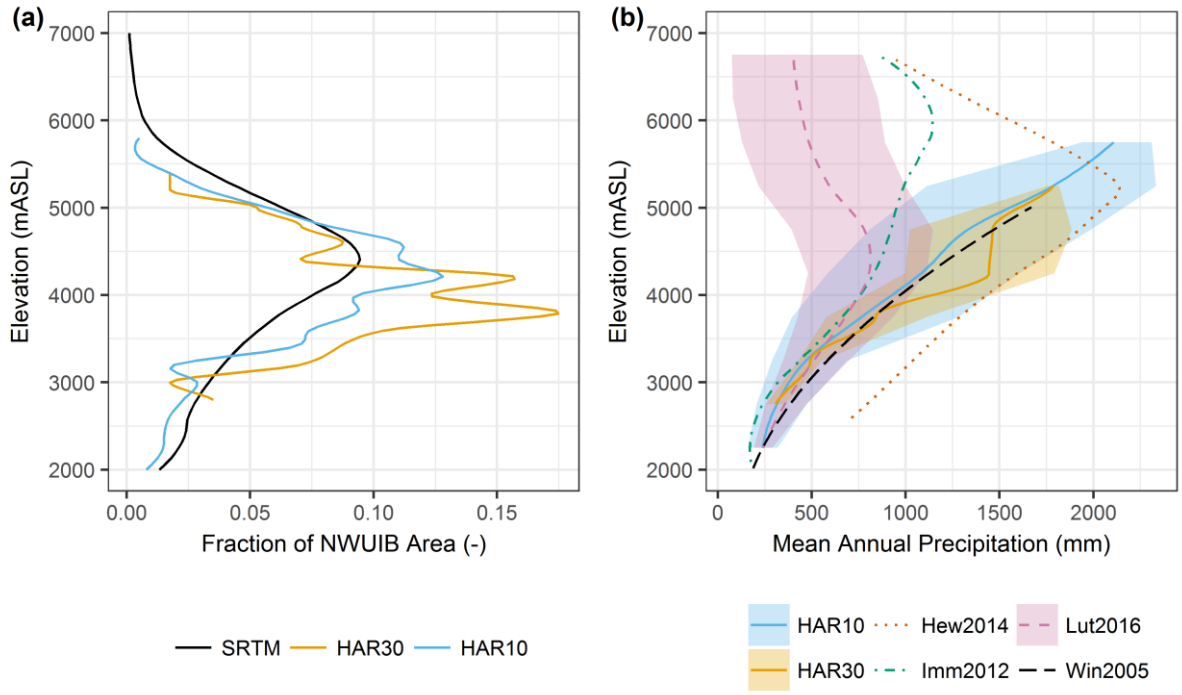


Figure A.7 – Comparison of mean annual vertical precipitation profiles between the HAR and other studies. Hypsometry for the NWUIB is given in (a), which compares the Shuttle Radar Topography Mission (SRTM) reference with HAR10 and HAR30. Vertical precipitation profiles are given in (b). Mean and ranges (10^{th} to 90^{th} percentiles) of the spatial climatology are shown for the HAR and Lutz et al. (2016), with the latter profile inferred from the data made available with their study. HAR and Lutz et al. (2016) profiles are calculated for the NWUIB, while other profiles were derived for different sub-domains of the NWUIB. Hew2014 is Hewitt (2014), Lut2016 is Lutz et al. (2016), Imm2012 is Immerzeel et al. (2012), and Win2005 is Winiger et al. (2005).

We reproduce the NWUIB hypsometry in Figure A.7a for reference. The vertical precipitation profiles are shown in Figure A.7b, which suggests that there is generally agreement in precipitation-elevation relationships up to around 3500 mASL. Above this elevation, the profiles differ significantly, which is at least partially a function of differences in methodology and the precise areas of the NWUIB considered in each analysis. Interestingly, the central tendency of the HAR10 vertical profile agrees quite closely with Winiger et al. (2005)'s profile, the formulation of which has been the basis for a number of other studies of UIB climate and hydrology (Bocchiola et al. 2011; Soncini et al. 2014; Reggiani et al. 2016). HAR30 shows a change in its precipitation-elevation relationship at the sharp changes in its hypsometric profile at around 4500 mASL, but its profile is more uncertain due to the smaller number of grid cells.

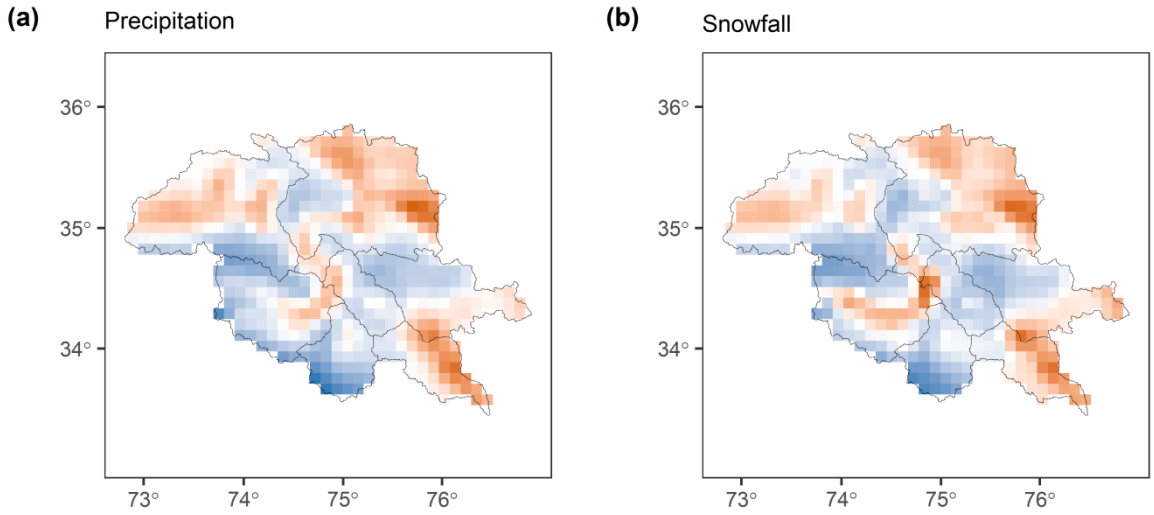


Figure A.8 – Comparison of standardised residuals after removing elevation signal for total annual (a) precipitation and (b) snowfall. Colour scale is as per Figure 4 in the main text.

The two profiles which provide an indication of spread in the relationship, HAR10 and Lutz et al. (2016), suggest that a substantial degree of spatial variation exists in precipitation-elevation scaling across this domain. This is likely to be very hydrologically significant, affecting precipitation phase simulations, runoff timing and glacier accumulation. It is also of note that the Lutz et al. (2016) and Immerzeel et al. (2012) profiles imply lower precipitation overall, particularly when considered in relation to the hypsometric profile.

In addition, we reproduce Figure 3.4c to compare the differences between using total annual precipitation and snowfall in Figure A.8. This shows that the patterns are very similar, owing to the high correlation between total precipitation and snowfall in the HAR. This is consistent with a higher fraction of annual precipitation falling in colder months and at higher elevations in the NWUIB. This also adds weight to the notion that a comparison of anomalies in precipitation and peak SWE derived from MODIS is possible at sub-regional scales, although at finer scales the influences of snow redistribution by wind and avalanches will reduce the correlation.

A.3 Temperature

A.3.1 Variability

In terms of variability, plotting the annual cycle of standard deviations of monthly temperature indicates good overall HAR performance (Figure A.9). The main difference is

that the HAR shows slightly higher inter-annual variability in summer and autumn. The observed peak in variability in May also happens a month later in the HAR, while the HAR has an additional peak in September/October that is less pronounced in observations.

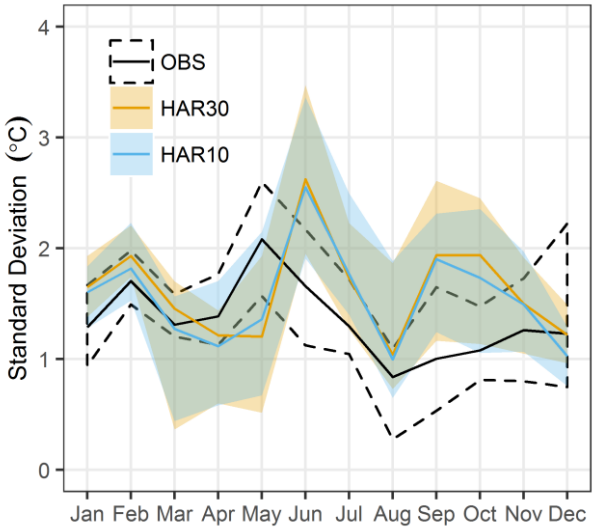


Figure A.9 – Comparison of the HAR and observed monthly standard deviation. Median and ranges calculated across all stations are shown.

A.3.2 Lapse Rates

We apply linear regression to estimate near-surface air temperature lapse rates for both the HAR and observations. We use the NWUIB domain in our estimation for the HAR and all available stations for the observed lapse rates. From Figure A.10a, we can see that the HAR and observed lapse rates are in good general agreement. The HAR exhibits higher lapse rates in summer, implying a faster rate of temperature decrease with elevation in the HAR. Indeed, there may be a connection here with the reduction in temperature bias in summer months (Figure 3.6a). We also note that there is closer agreement between HAR10 and HAR30 in summer months.

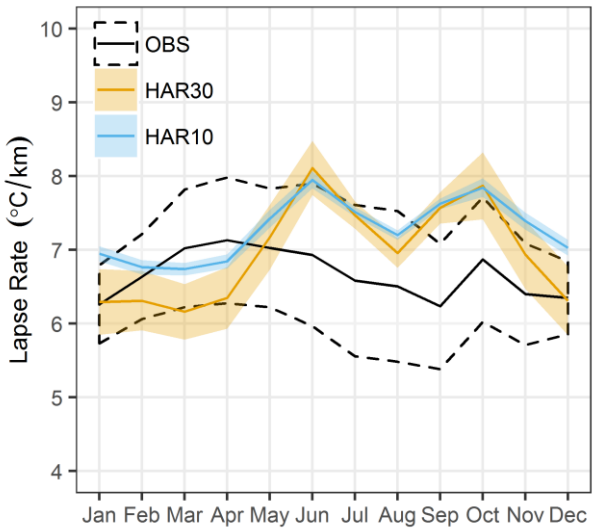


Figure A.10 – Comparison of lapse rates based on monthly mean temperatures determined through linear regression for the HAR (NWUIB domain) and observations. The ranges show 95% confidence intervals for the estimated lapse rates.

A.4 Humidity

Figure A.11 shows winter and summer vertical profiles for relative humidity (as opposed to specific humidity, given in Section 3.4.3). The most striking feature here is that the HAR shows a clear increase in relative humidity with elevation in winter, whereas the observations contain very little elevation dependence. In contrast, the nonlinear profile of relative humidity increase with elevation in summer in the HAR appears consistent with the available observations. As HAR temperature lapse rates are reasonable in winter, the difference could be attributable to a greater influence of cold temperature bias on the relative lowering of saturation specific humidity at higher elevations. Lesser absolute temperature biases in summer may help to explain the greater conformity in the HAR's relative humidity profiles at this time of year, although the lower peak in the specific humidity cycle may also contribute. HAR10 performance is better than the HAR30 for relative humidity in both seasons, which accords with its typically lower absolute temperature bias throughout the year.

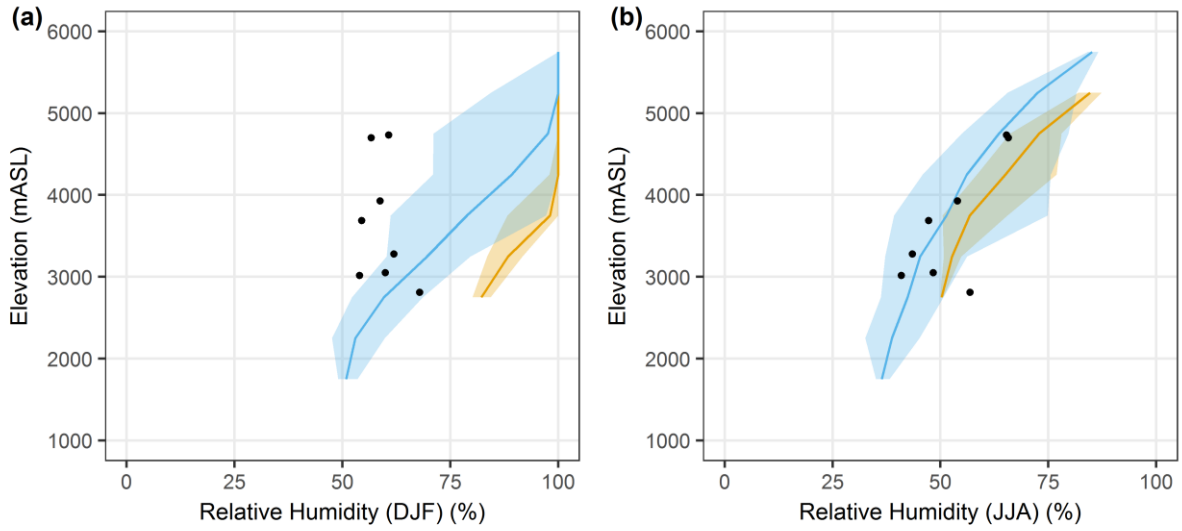


Figure A.11 – Comparison of observed and HAR vertical profiles of relative humidity for (a) winter (DJF) and (b) summer (JJA). HAR profiles show the elevation band means and ranges.

A.5 Incoming Shortwave Radiation

To support the finding of a clear-sky bias that peaks in summer, Figure A.12 shows incoming shortwave radiation time series for additional stations and years. This confirms that the HAR accurately simulates incoming shortwave radiation under clear-sky conditions at all available station locations. It also demonstrates that incoming shortwave variability is underestimated by the HAR in summer in different years and at all stations, particularly Askole and Urdukas. Lower data availability for Concordia makes it harder to be definitive for this station, but there does seem to be overestimation by the HAR in the summer months in 2012. This supports the conclusion drawn in Section 3.4.4, namely that the HAR exhibits underestimation of cloudiness or cloud reflection effects.

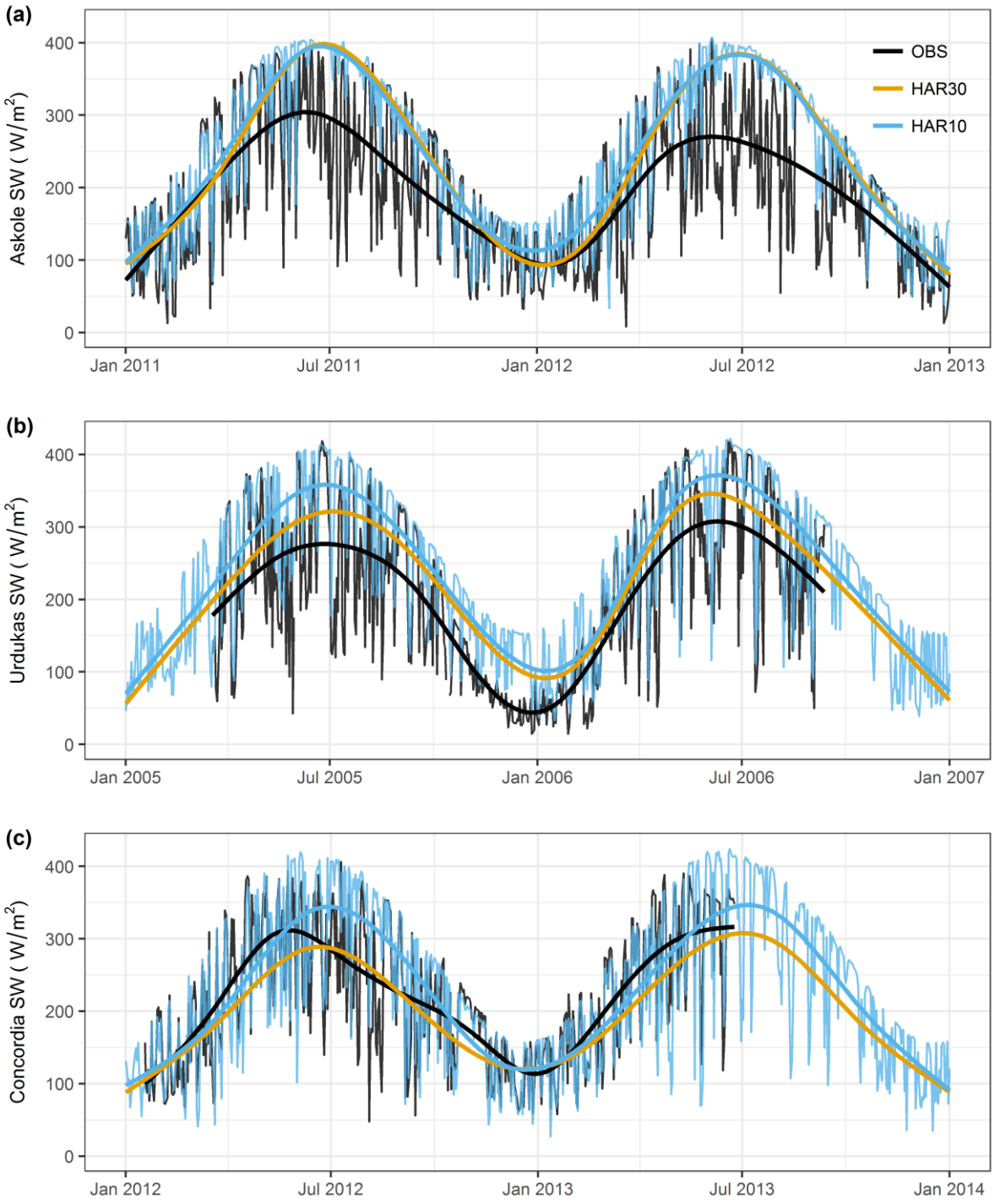


Figure A.12 – Time series of daily incoming shortwave radiation (SW) for (a) Askole, (b) Urdukas and (c) Concordia AWSs for both observations and corresponding HAR cells. Bold lines show 28-day moving average series.

A.6 Wind Speed

Due to its importance for turbulent heat fluxes between the atmosphere and surface, we made a preliminary comparison of HAR wind speeds with observations from the EvK2CNR

stations. We used a logarithmic profile (assuming stability) to adjust the HAR wind speeds from their height above the surface (10 m) to the observation heights. This shows that the HAR cells have higher wind speeds compared with observations – between 2 and 5 m/s higher on average – as well as higher inter-annual variability. This is likely due to the significant differences in scale between point measurements and HAR grid cells, a particularly pertinent issue here, given the substantial variability of wind at multiple scales. However, the HAR and observed coefficients of variation are relatively similar (around 0.3), suggesting reasonable HAR performance in relative terms.

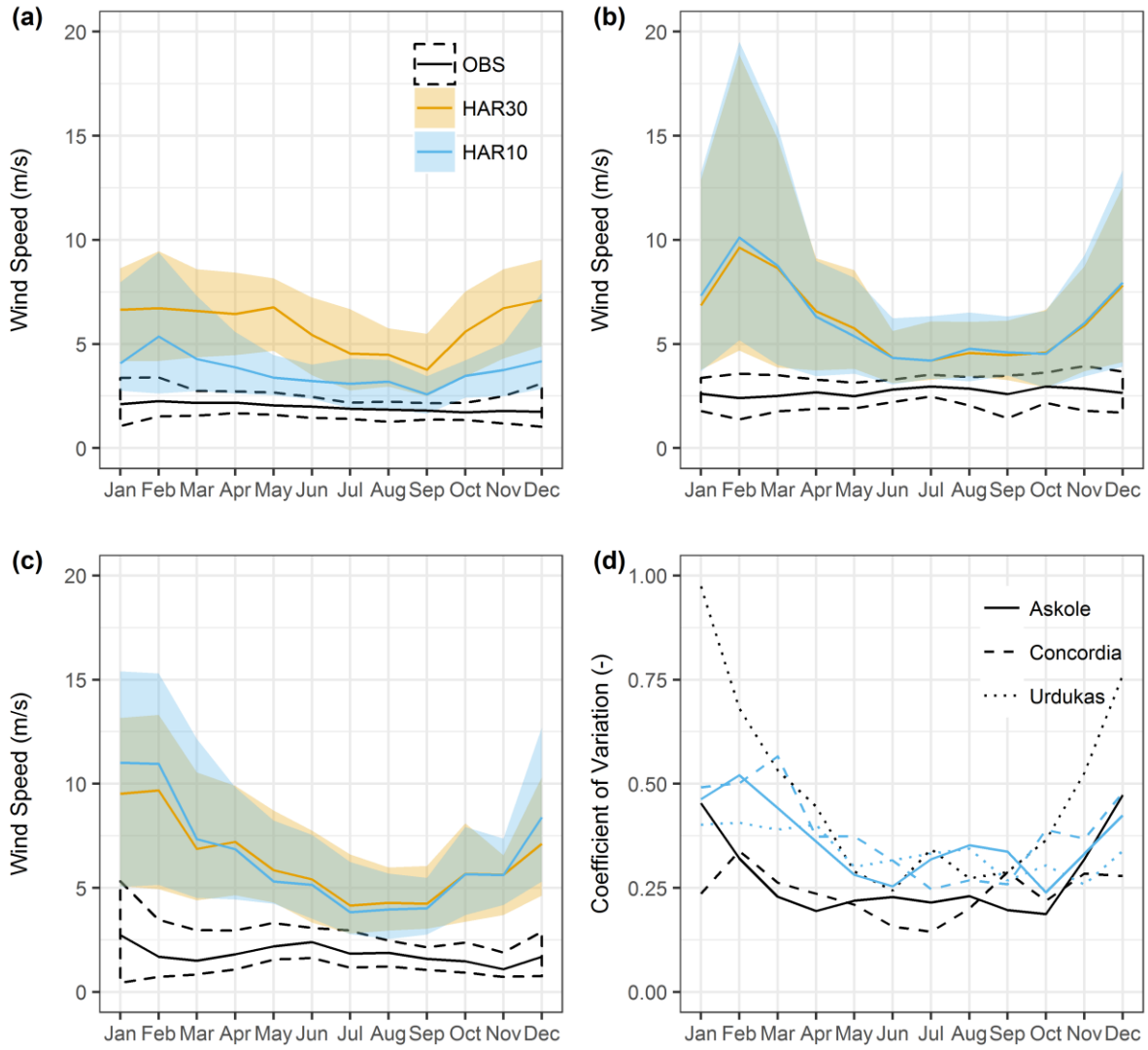


Figure A.13 - Wind speed climatology comparison. Annual cycles of mean monthly wind speeds are shown for (a) Askole, (b) Concordia and (c) Urdukas EvK2CNR stations. Comparison of monthly coefficients of variation for the same stations is given in (d).

A.7 Cloud, Albedo and Temperature Correlations

In Table A.2 we provide a summary of the significance of correlations presented in Figure 3.12.

	Cloud Cover Fraction vs Temperature				**Albedo vs Temperature**			
	HAR10		Observations		HAR10		Observations	
Month	P<0.1	P<0.05	P<0.1	P<0.05	P<0.1	P<0.05	P<0.1	P<0.05
Jan	2	0	0	0	2	0	3	2
Feb	1	0	0	0	0	0	0	0
Mar	0	0	2	1	0	0	2	1
Apr	4	2	0	0	4	3	4	1
May	0	0	3	1	7	6	3	0
Jun	6	4	2	1	10	10	1	1
Jul	0	0	7	5	9	7	0	0
Aug	0	0	2	1	1	0	3	1
Sep	0	0	2	2	10	10	1	0
Oct	0	0	0	0	10	10	1	0
Nov	0	0	1	0	2	1	2	1
Dec	4	4	6	2	2	1	0	0

Table A.2 – Number of stations with significant correlations for each month based on threshold P values of 0.1 and 0.05. The total number of stations used in the correlations is 10.

Appendix B

Supporting Information for Chapter 4

B.1 MODIS Land Surface Temperature (LST) Evaluation

Here we provide additional validation of MOD11A1 Collection 6 land surface temperature (LST). The LST values are based on the average of the 9 pixels surrounding a station location and are corrected for elevation differences using local lapse rates (estimated from linear regression). Beginning with a direct comparison at the Concordia site, one of very few locations with outgoing longwave radiation measurements with which to estimate LST, we can see that the MODIS data show good correspondence with observations overall (Figure B.1 and Figure B.2).

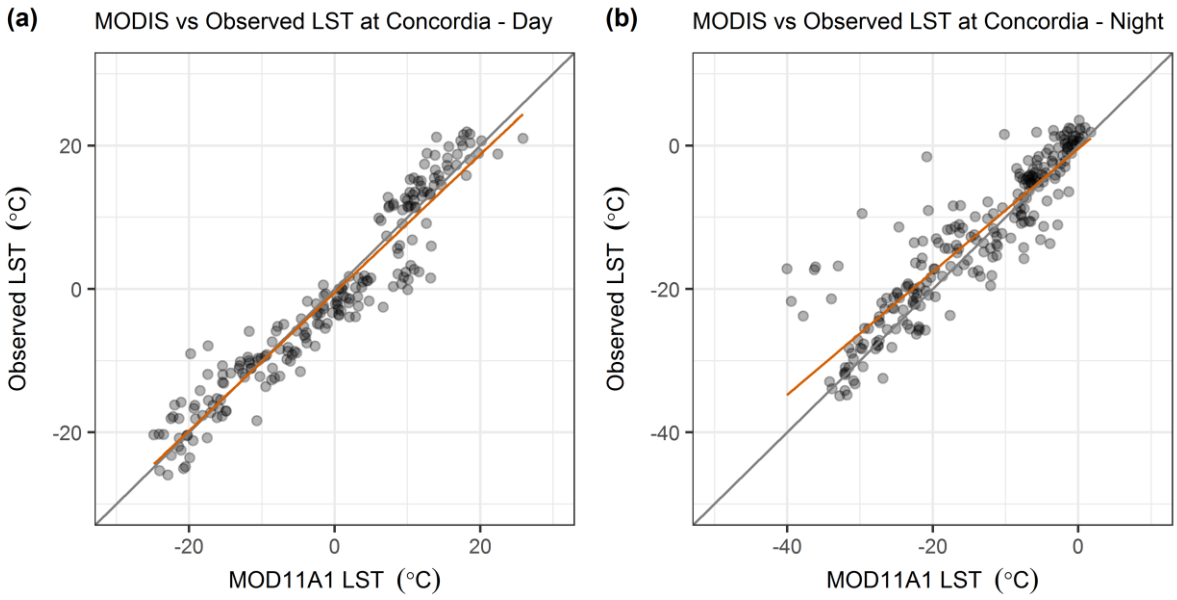


Figure B.1 – Comparison of LST from observations (of outgoing longwave radiation) and MODIS remote sensing at the Concordia site.

Season	Bias (°C)		RMSE (°C)	
	Night	Day	Night	Day
Annual	-1.6	0.5	5.0	3.7
DJF	-3.5	-1.4	3.4	7.2
MAM	-0.2	2.1	3.6	4.7
JJA	-0.7	1.4	5.0	3.7
SON	-2.8	-1.4	2.2	4.3

Table B.1 – Summary statistics for MOD11A1 performance at Concordia site.

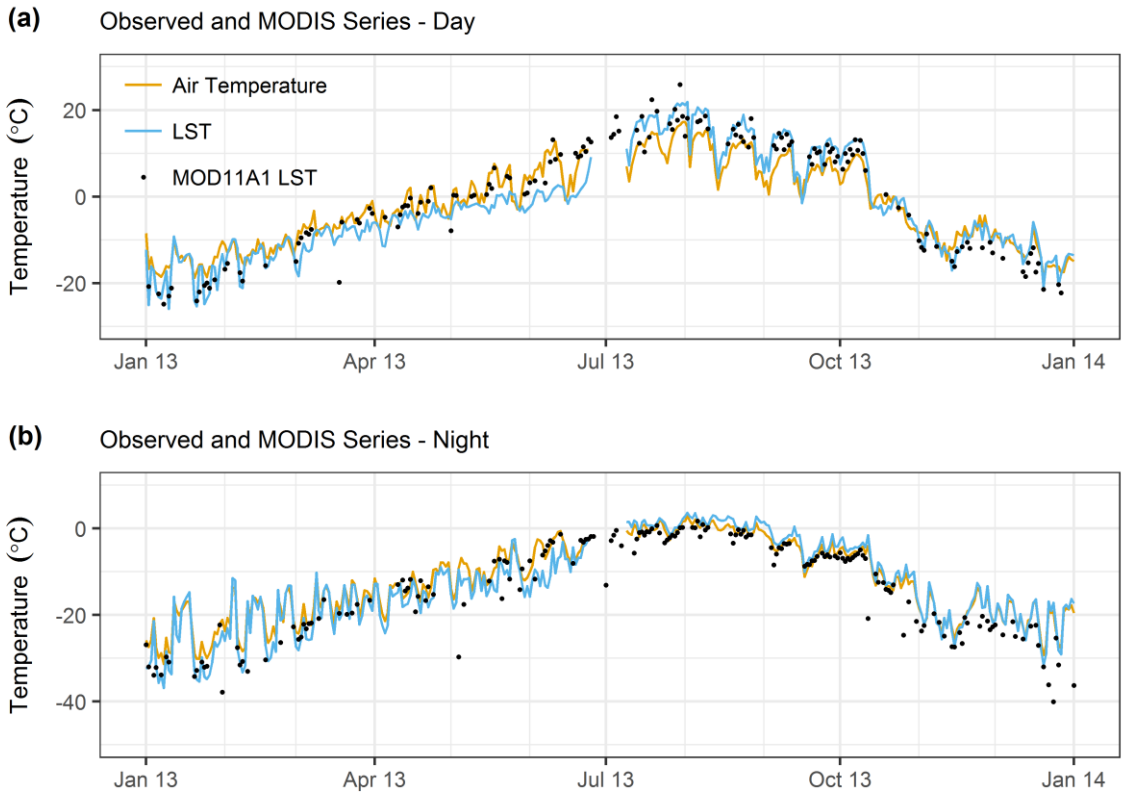


Figure B.2 – Time series comparison of observed LST and air temperature with MODIS LST for the Concordia site.

Summary statistics in Table B.1 confirm that bias is slightly larger for individual seasons than at the annual timescale, but overall it is low. Importantly, this indicates that MOD11A1 is

accurate enough to estimate climatological LST to within 2-3°C, depending on season. Therefore, modelled LST that departs from MODIS retrievals by notably more than this is likely to be in error.

We also examine the relationship between LST and air temperature. For the Concordia site, where both quantities are available, Figure B.3 shows that there is strong similarity between LST and air temperatures. This is particularly the case at night-time, where the relationship is almost 1:1 and scatter is much reduced, although there is some divergence at the lowest temperatures. The relationship in the day-time is more complicated and variable, showing a discontinuity around 0°C. This likely indicates abrupt changes as a result of snow cover disappearance.

Unfortunately observed LST is not available at other climate stations. As such, we compare observed air temperature with MODIS LST at these locations in Figure B.4 and Figure B.5. This shows that there is also a close, often approximately 1:1 relationship between LST and air temperature at night-time at other station locations, strengthening the findings from the Concordia site. The day-time case generally shows that air temperatures are warmer than LST in cold conditions, but that the opposite is true in warmer conditions. Much more scatter is present relative to the night-time comparison.

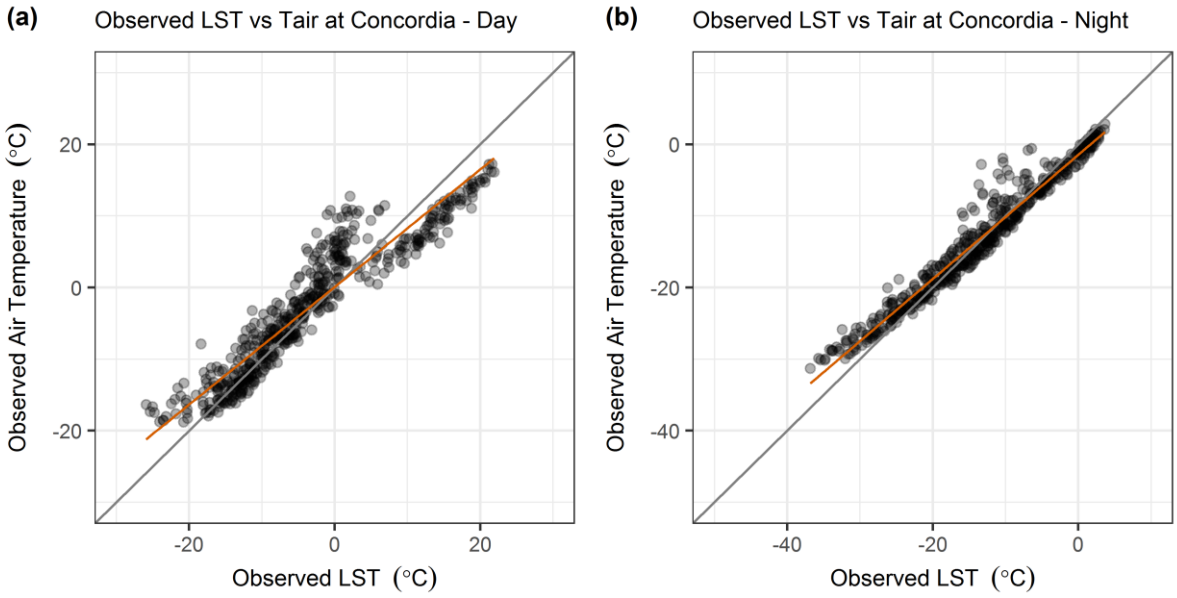


Figure B.3 – Comparison of observed LST (calculated from outgoing longwave radiation) and air temperature at the Concordia site.

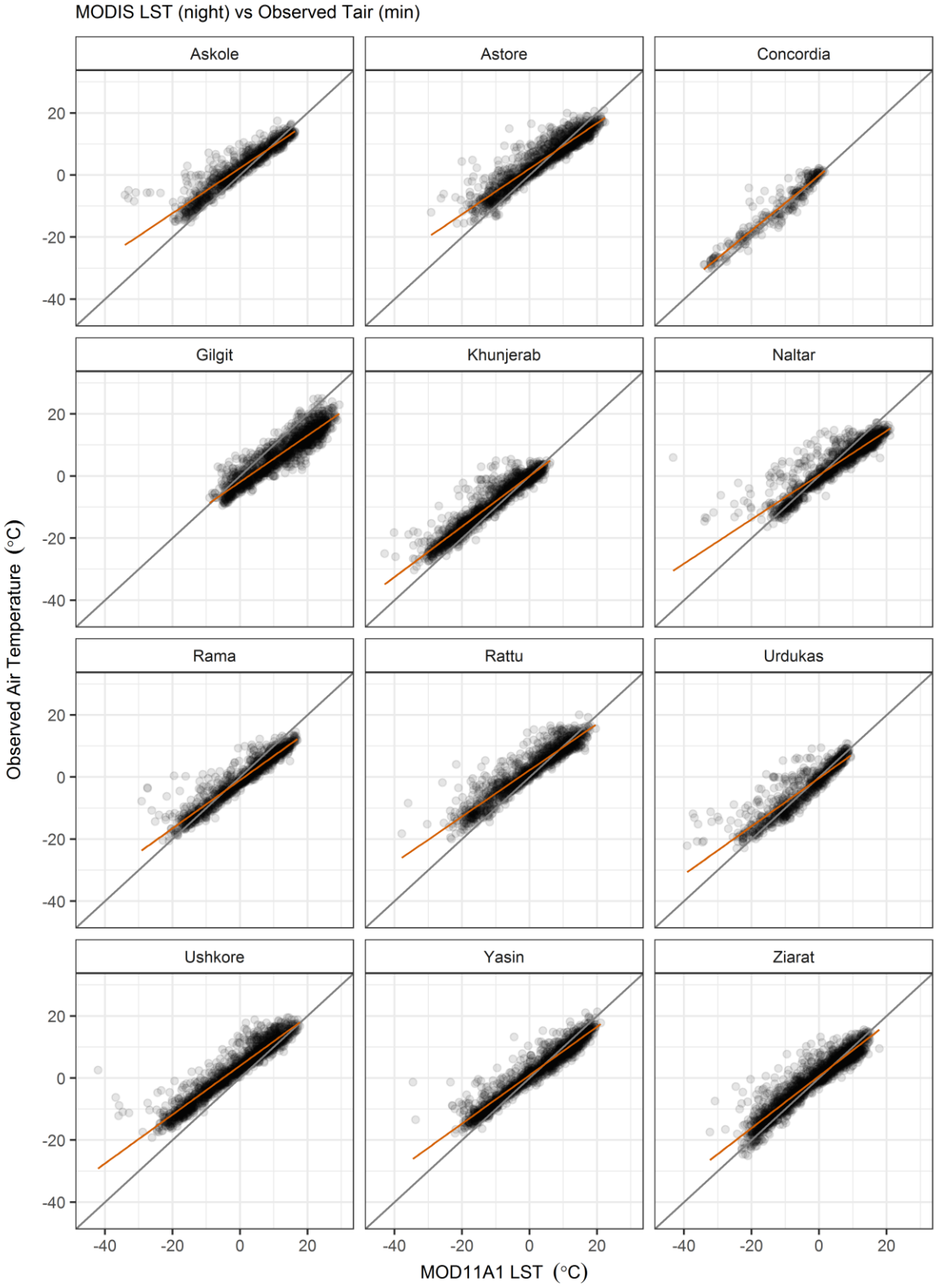


Figure B.4 – Comparison of observed night-time / minimum air temperatures with MODIS LST at observation sites that do not measure outgoing longwave radiation.

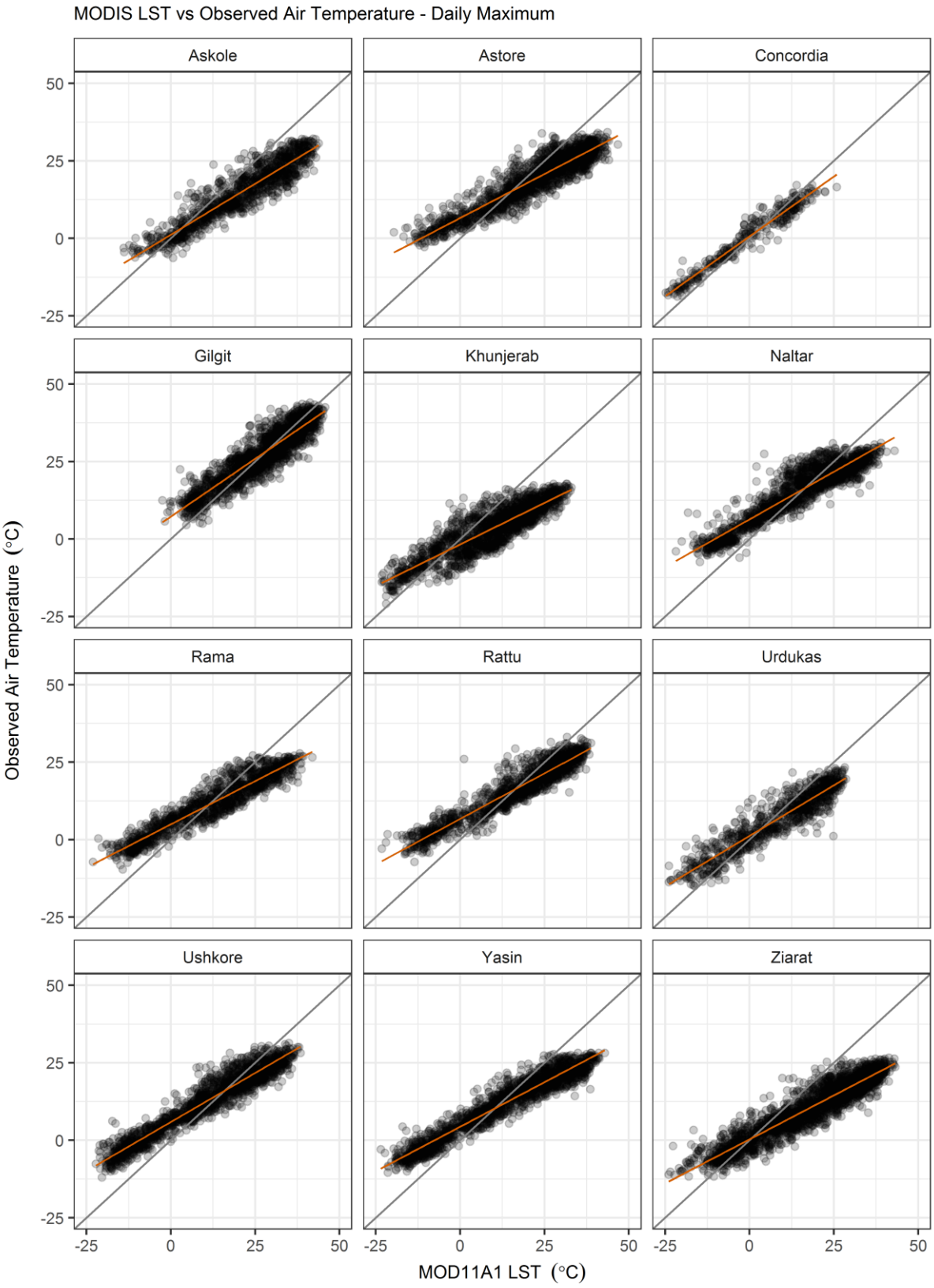


Figure B.5 – Comparison of observed day-time / maximum air temperatures with MODIS LST at observation sites that do not measure outgoing longwave radiation.

B.2 Point-Scale Simulations at Concordia

The Concordia site is situated at 4700 mASL on supra-glacial debris in the upper zone ablation of the Baltoro glacier in the Shigar sub-basin of the Upper Indus Basin (UIB). This location is shown on Figure 4.1 in Chapter 4. Concordia is one of very few sites in the region recording the climate variables required to simulate (and partially evaluate) the energy balance in FSM. No such sites are present in the Astore sub-basin. As such, we use simulations at Concordia for verification of the FSM code and for selected comparisons with the distributed simulations. The key results from this are given in Section B.5.

We obtained quality-controlled data for Concordia from EvK2CNR, as well as additional raw data that we checked for consistency and outliers. The available record period for Concordia is short, such that we conducted two-year simulations, cycling the climate inputs to equilibrate the substrate state variables. We filled small gaps in the hourly time series using simple interpolation. The data include measurements of outgoing shortwave and longwave radiation, which we use to assist in model evaluation. However, precipitation is not reliably recorded at Concordia. To overcome this, we use data from the High Asia Refined Analysis (HAR) (Maussion et al., 2014).

FSM is based on the presumption of soil layers beneath any snow present. This is not the case at the Concordia site. However, we use the results of point-scale simulations here to support particular findings from the distributed modelling in the Astore catchment – specifically conclusions on albedo evolution and stability effects on land surface temperature that depend less heavily on substrate properties. While this modelling therefore represents a notable simplification of the reality of the site, we consider it useful to further understand FSM behaviour in this data-sparse context.

B.3 Ensemble Evaluation for Alternative Input Strategies

In this section we provide evaluation of the distributed FSM simulations for the Astore catchment using the two alternative input strategies described in Section 4.3.2. One strategy is essentially HAR-based with no bias correction (FSM2), whereas the other strategy uses local observations as far as possible (FSM3). From Figure B.6 and Figure B.7 we can see that simulated runoff in both strategies is reasonable. FSM2 shows a generally similar form to FSM1 reported in the main text, whereas FSM3 shows a little more difference. In particular, ensemble spread using FSM3 is larger and the coherence of the ensemble sub-groups

decreases faster. At the ensemble level, runoff is also a little earlier in FSM3, leading to lower peak runoff rates in the middle of the season, although some ensemble members actually tend to respond quite slowly.

Similar comparisons are shown for snow-covered area (SCA) in Figure B.8 and Figure B.9. This confirms that SCA is reasonably well simulated when using both FSM2 and FSM3 input strategies. Larger spread is evident for FSM3, consistent with the runoff responses described above. In combination with adequate runoff performance, these results suggest that the alternative input strategies are sufficient to see if key findings based on FSM1 are also evident. This provides a partial treatment of uncertainty, as discussed in Section 4.3.2.

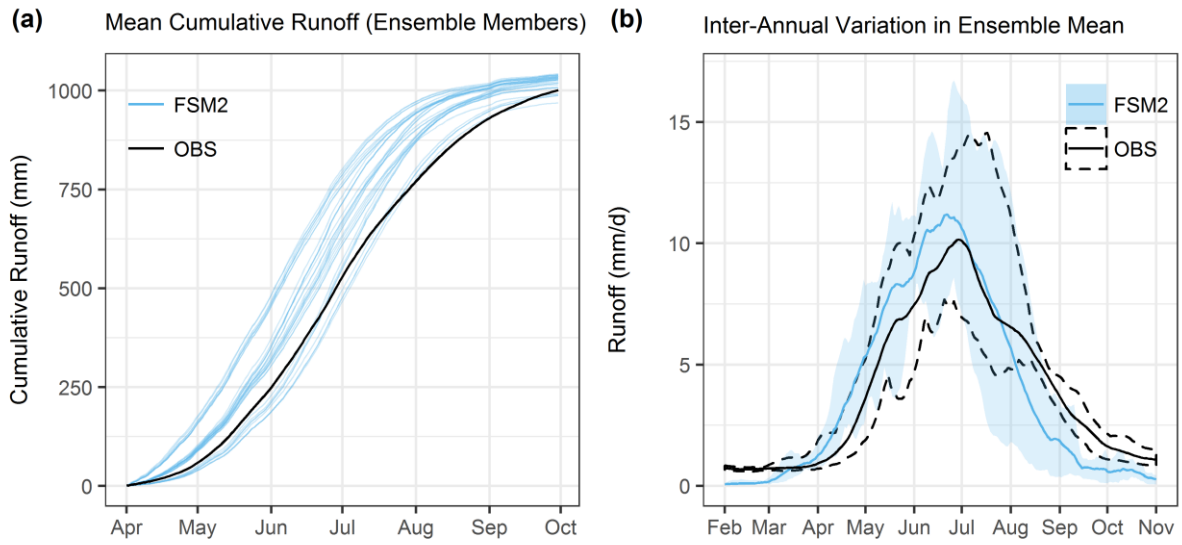


Figure B.6 – Comparison of modelled snowmelt runoff with observed runoff for the FSM2 input strategy. Mean annual cumulative runoff for the high-flow season for each of the 32 ensemble members is given in (a), along with observations. Inter-annual variation in the ensemble mean and observed runoff is shown in (b), where the solid lines denote the mean and the shaded/dashed areas indicate the inter-annual range. Model results are are shown in blue, with observations in black.

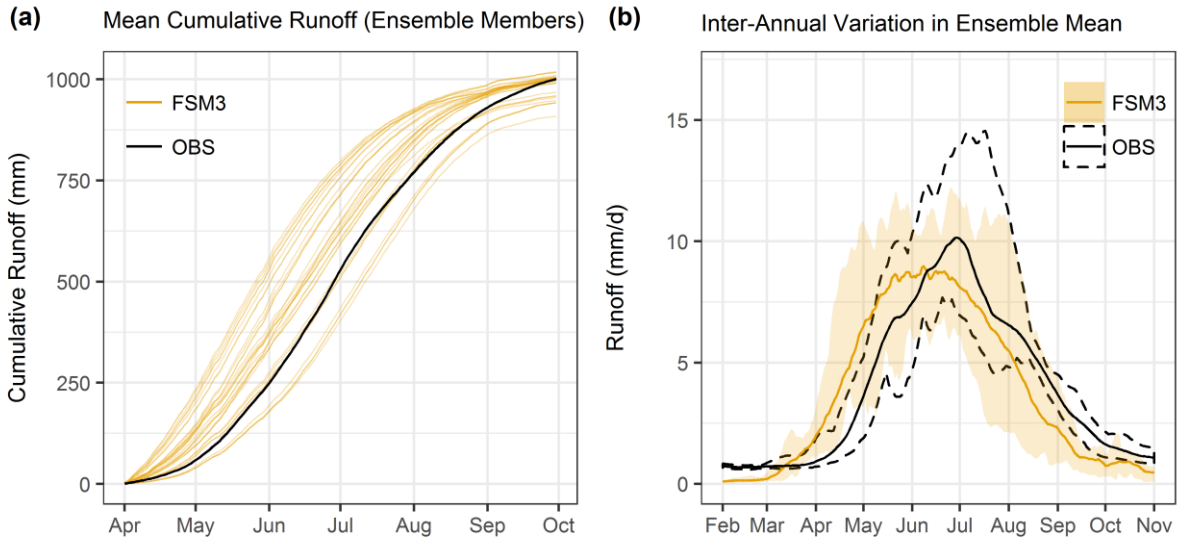


Figure B.7 – As Figure B.6 but for FSM3 input strategy (modelled series in orange).

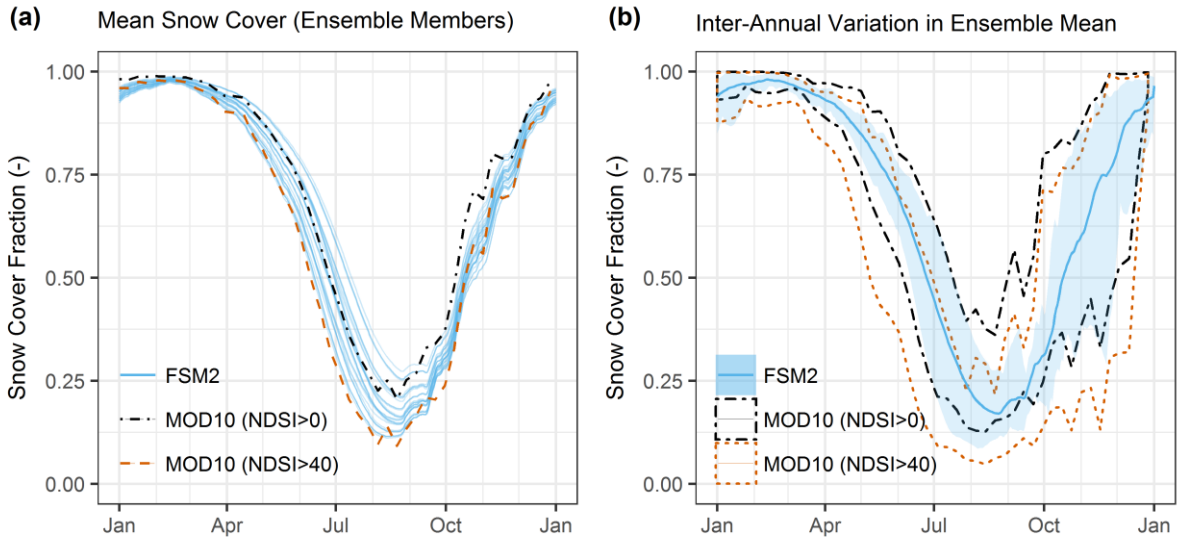


Figure B.8 – Comparison of modelled catchment snow cover with MODIS remote sensing for FSM2 inputs. The mean annual cycle of snow cover for each ensemble member is displayed in (a), along with MOD10A1 snow cover. Inter-annual variations in the ensemble mean and MOD10A1 snow cover is plotted in (b) as the dashed/shaded areas.

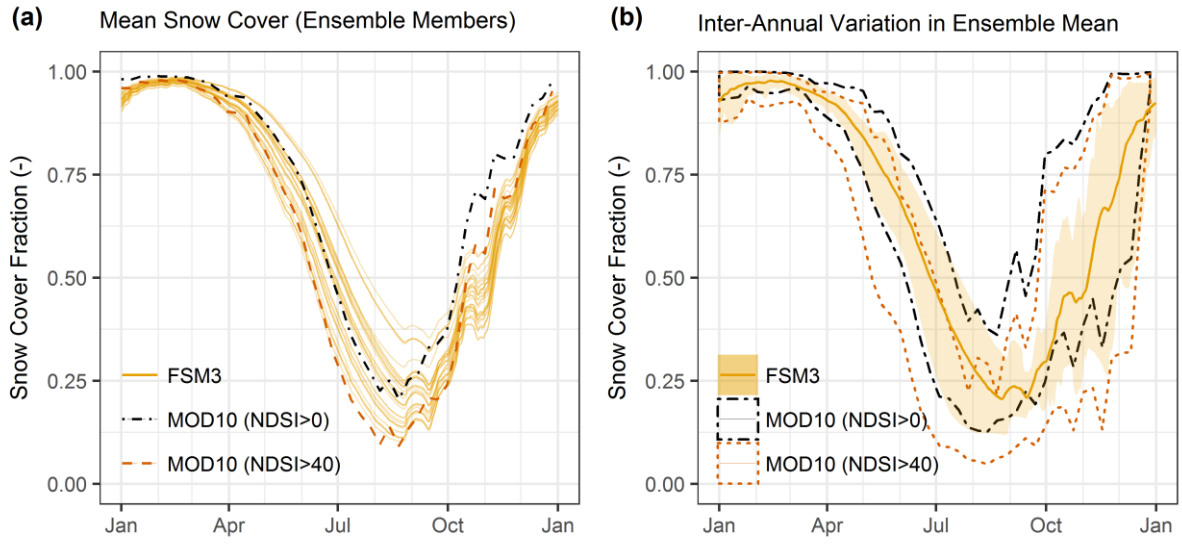


Figure B.9 – As Figure B.8 but for FSM3 inputs (model series in orange).

B.4 Process Influence on Runoff Response

In Figure B.10 we confirm that the thermal conductivity option plays a comparatively minor role in shaping ensemble spread in runoff.

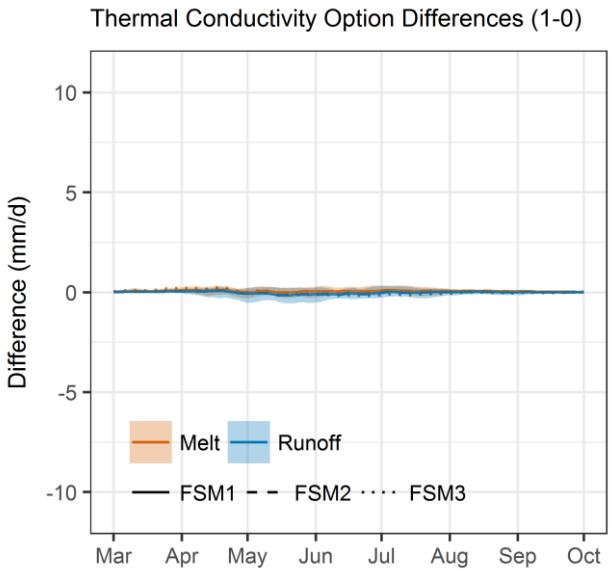


Figure B.10 – Differences in melt and runoff arising from thermal conductivity option choices (as Figure 4 in main text).

Figure B.11 additionally shows the average effects of each process option choice on timing. This approximately indicates differences in runoff onset and cessation, by quantifying how

many more/less days are needed to exceed the 10th and 90th percentiles of cumulative runoff each year when using different process options. The main point from this is that the density option can exert a notable effect on runoff timing in some years, especially in the later part of the melt season. Possible mechanisms for this include FSM's use of snow depth to adjust albedo for patchy cover, which confirms the importance of subgrid variability in these conditions (Clark et al., 2011). Other possibilities include influences on porosity and so liquid water holding capacity and related processes. While we focus mainly on albedo, drainage and stability adjustment options in this study, the hydrological implications of density and thermal conductivity options should be explored further.

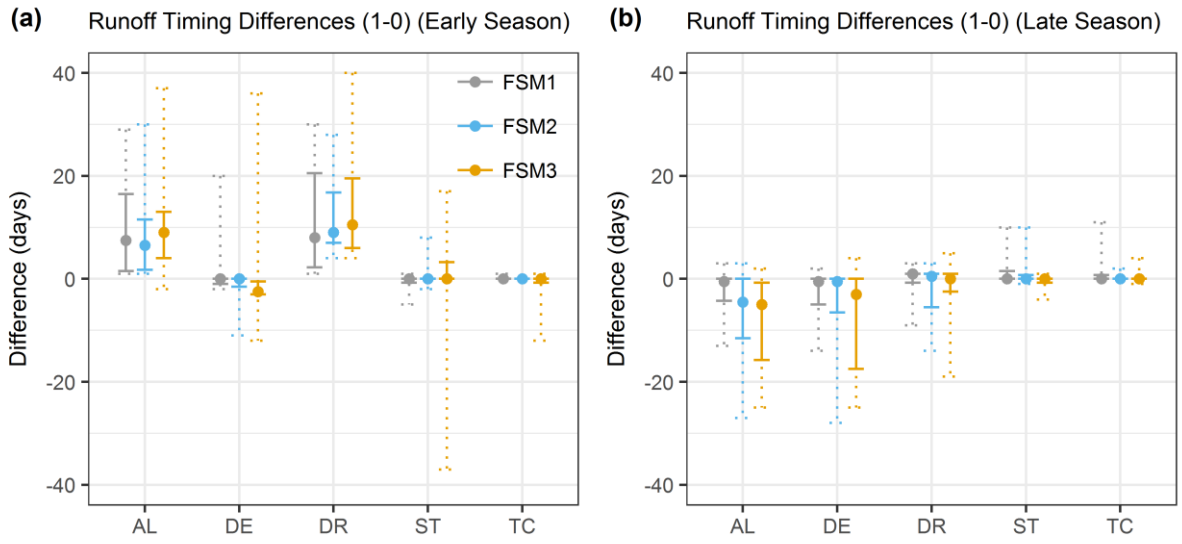


Figure B.11 – Differences in runoff timing according to model options. The differences are calculated by first averaging all ensemble members using options 0 and 1, separately for each of the 5 processes (AL – albedo, DE – density, DR – drainage, ST – stability, TC – thermal conductivity). The day of year at which the 10th (a) and 90th (b) percentiles of annual cumulative runoff are exceeded are then identified and differences calculated (option 1 minus option 0). Positive differences therefore indicate later timing when using option 1 for a given process and vice versa. The average (dot), inter-quartile range (solid lines) and range (dotted lines) for all years are plotted. Colours denote the different input strategies.

B.5 Albedo Evaluation

We now evaluate the point-scale simulations at Concordia with reference to albedo options, in order to support the findings in Section 4.4.3. Figure B.12a confirms that the prognostic albedo parameterisation is more realistic at the point-scale. This is based on observations of incoming and outgoing shortwave radiation, rather than remote sensing. Figure B.12b shows

that LST is well simulated at the site. This indicates that fundamental energy balance inaccuracies, for example due to input data uncertainties, are not the cause of poor performance with the diagnostic albedo parameterisation. This is likely to be the case even when the LST-albedo feedback induced by this parameterisation is taken into account.

Plotting albedo against LST in Figure B.13 shows that there is substantial scatter in their relationship, which depends at least partly (or indirectly) on the time of year. Similar patterns are evident when using observations from Concordia and MODIS retrievals (for pixels with NDSI greater than 70, i.e. substantial snow cover). This is further confirmation that a simple deterministic parameterisation of albedo based on LST is not the most appropriate representation in this context. The effects of fresh snowfall and subsequent albedo decay during dry periods are important and better simulated using a prognostic parameterisation.

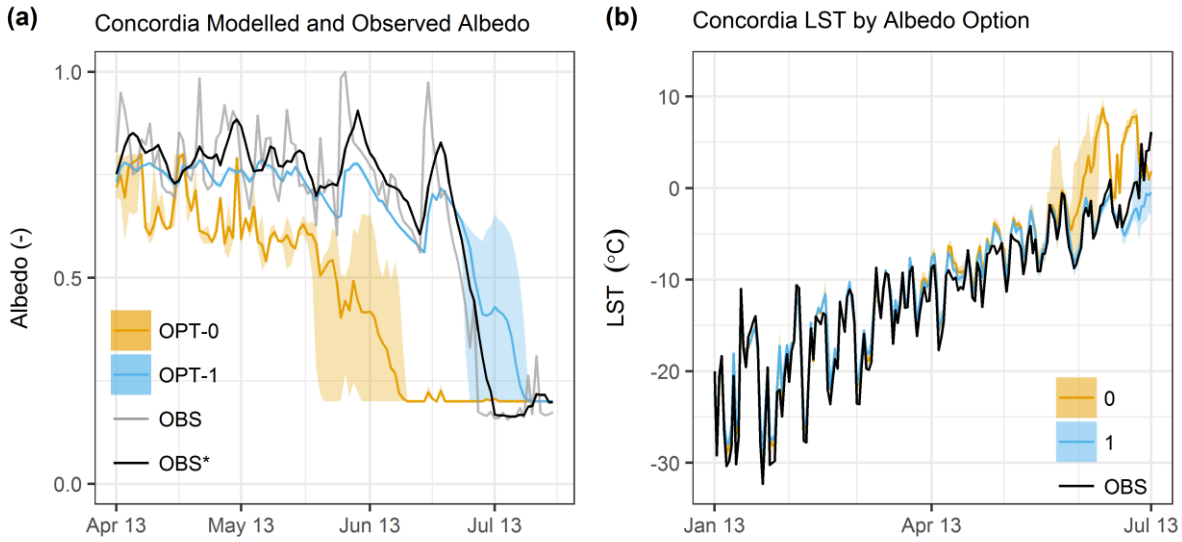


Figure B.12 – Modelled and observed daily mean albedo (a) and LST (b) at Concordia. The OBS* series in (a) corresponds with an 8-day moving average. The FSM ensemble is split according to members using option 0 (orange) and 1 (blue). The mean (line) and range (shading) of each group are shown. Daily observed albedo calculated from incoming and outgoing shortwave radiation measurements are shown in grey in (a), with the 8-day moving average shown in black (OBS*).

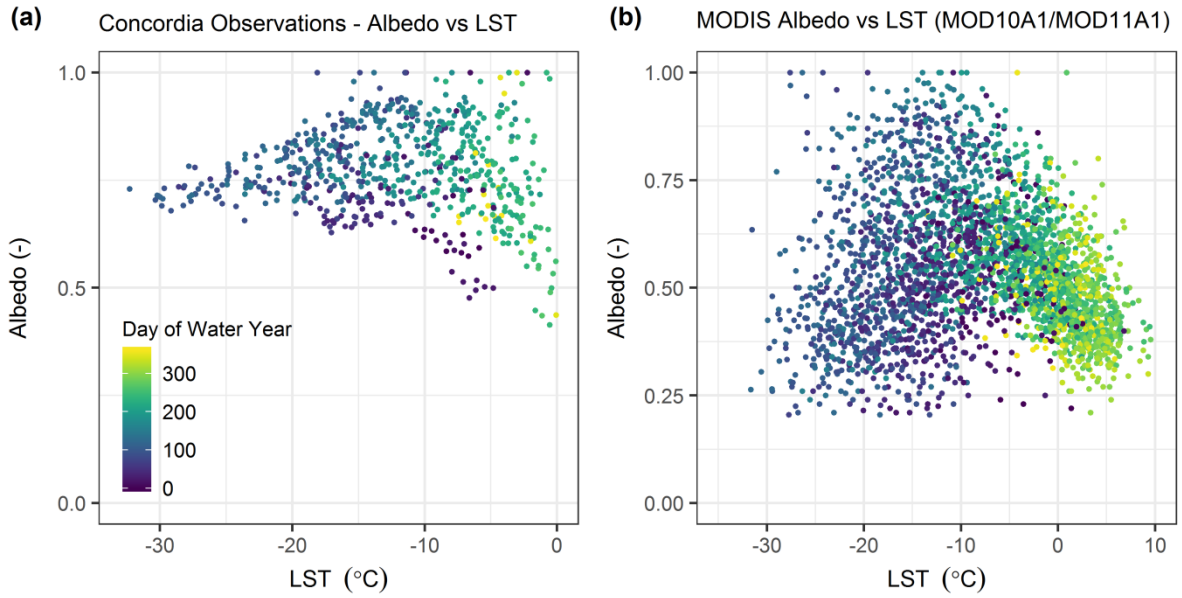


Figure B.13 – Relationship between LST and albedo based on observations at Concordia (a) and MODIS remote sensing (b). The dots are coloured by the day of the water year (defined as October to September).

B.6 Drainage Evaluation

With the drainage option on, liquid water is allowed to refreeze, leading to latent heat release. This maintains a higher snowpack temperature, which is conducive to higher melt rates. This is exemplified in Figure B.14, which shows how the drainage option choice becomes a primary control on snowpack temperature from the onset of melting. Following ripening, the ensemble spread in snowpack temperature for models applying the drainage option is very small and remains close to melting point, which also implies that the diurnal range is small. This is conducive to higher melt rates.

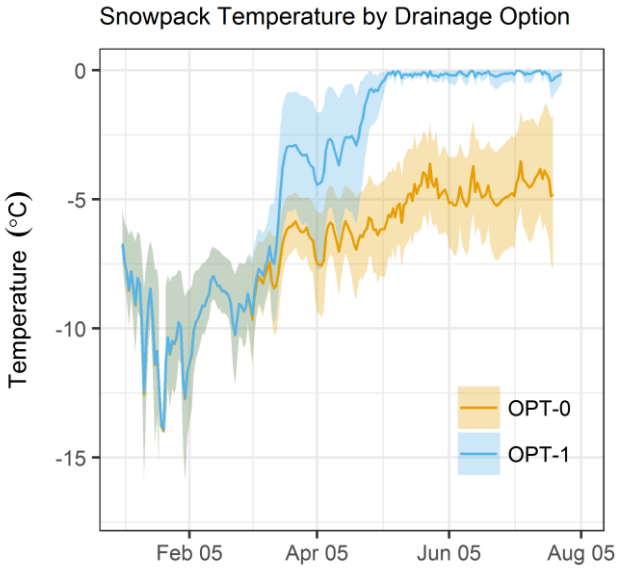


Figure B.14 – Comparison of snowpack mean (mass-weighted) temperatures for elevation band 3500-4000 mASL using each drainage option.

B.7 Process Interactions

Examples of cumulative runoff curves grouped by their major controls are given in Figure B.15. This confirms the general year-to-year consistency in ensemble structure discussed in Section 4.4.4. It also shows how the two intermediate groups exhibit differing degrees of equivalence and rank order in different years.

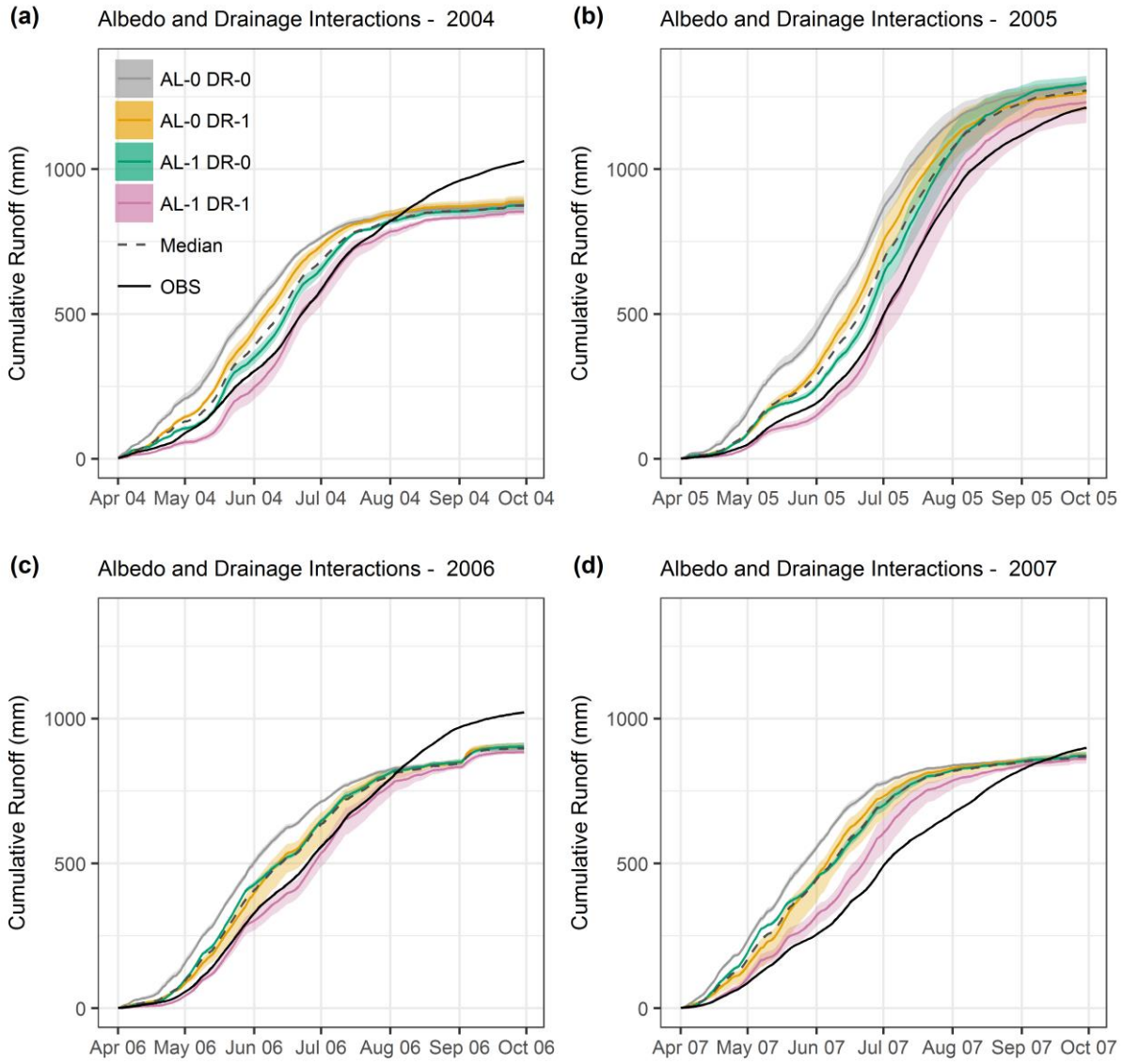


Figure B.15 – Examples of ensemble interactions / groupings for individual years (see Figure 10 in main text).

B.8 Performance Relationships

To expand on the relationships between model performance for different ensemble groups, as discussed in Section 4.4.5 in the main text, we show scatterplot matrices in Figure B.16 and Figure B.17 for runoff and SCA RMSE in individual years. This indicates whether performance for a given year is similar in different model groups, or whether more complicated relationships exist. From the matrices we can see that a range of relationships are present. These include the Pareto fronts discussed in Section 4.4.5, which primarily develop for better-performing configurations. In other cases there are simple, positive linear relationships, indicating that performance for a given year is good or bad in both

configurations in question. There are also examples of very limited relationships in performance, as well as multiple linear relationships for the same pairs.

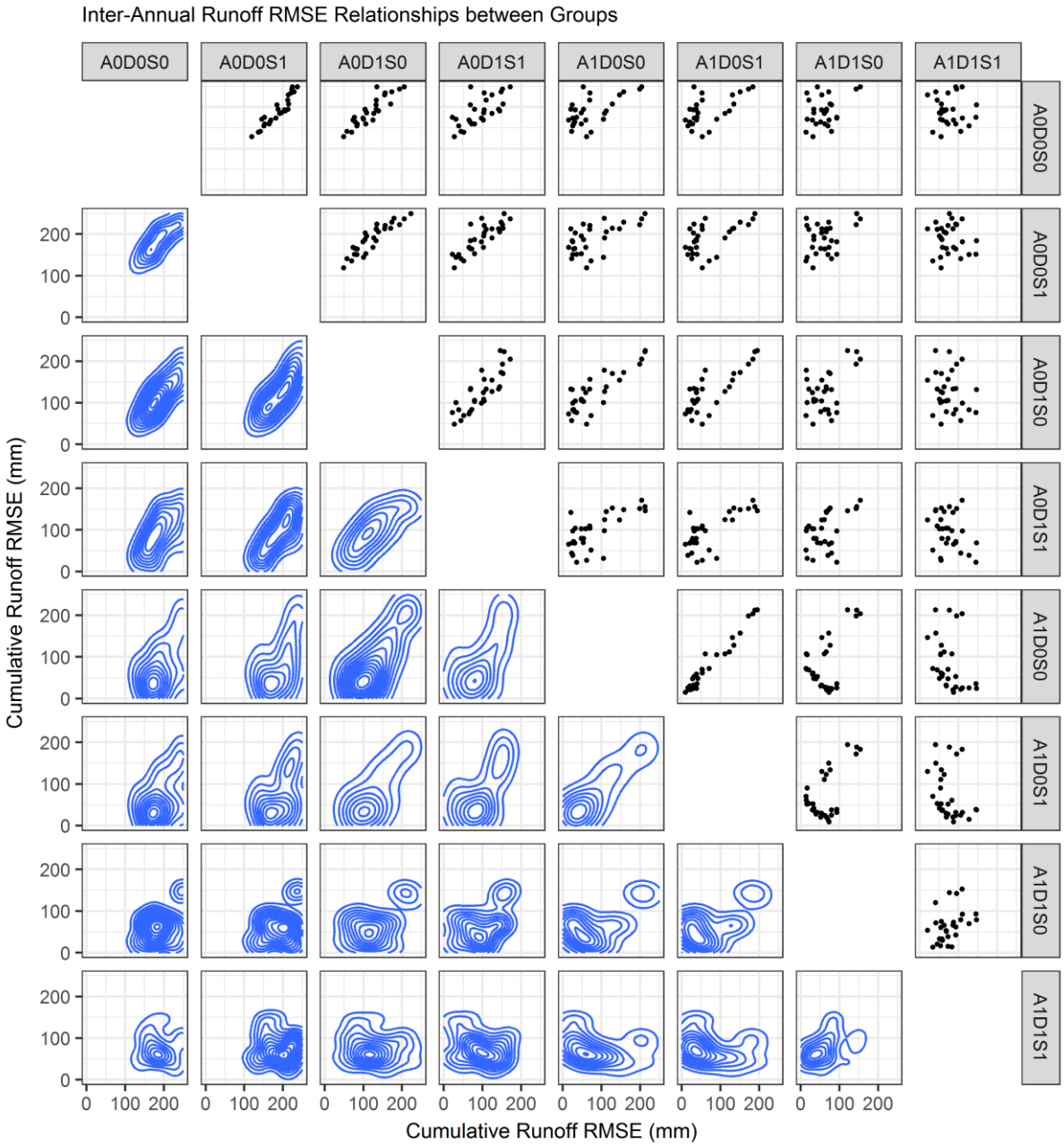


Figure B.16 – Scatterplot matrix comparing cumulative runoff RMSE for ensemble groups using different configurations of albedo (A), drainage (D) and stability adjustment (S) options. Scatterplots are shown in the upper right and density estimations in the lower left.

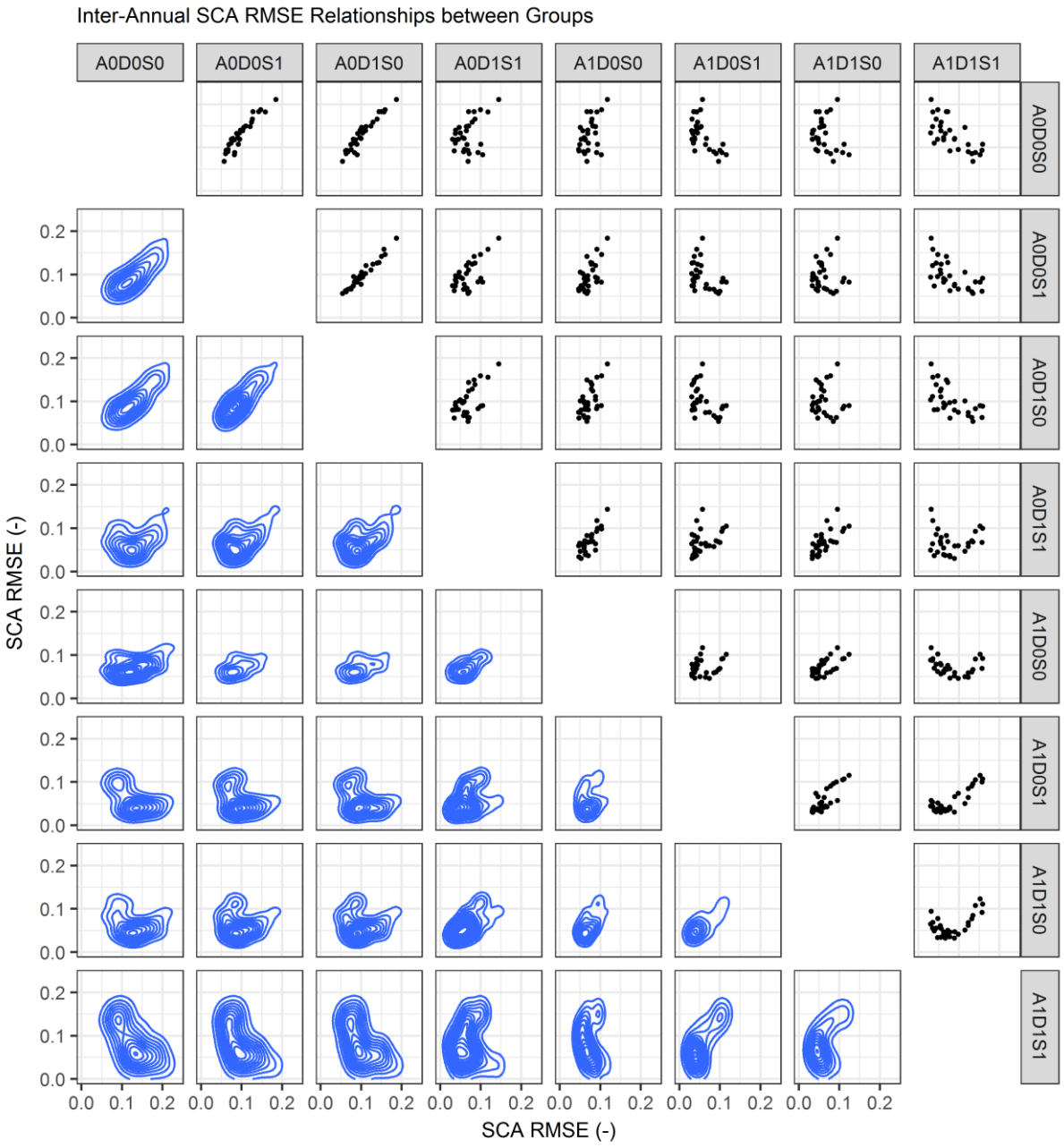


Figure B.17 – As Figure B.16 but for SCA.

Appendix C

Supporting Information for Chapter 6

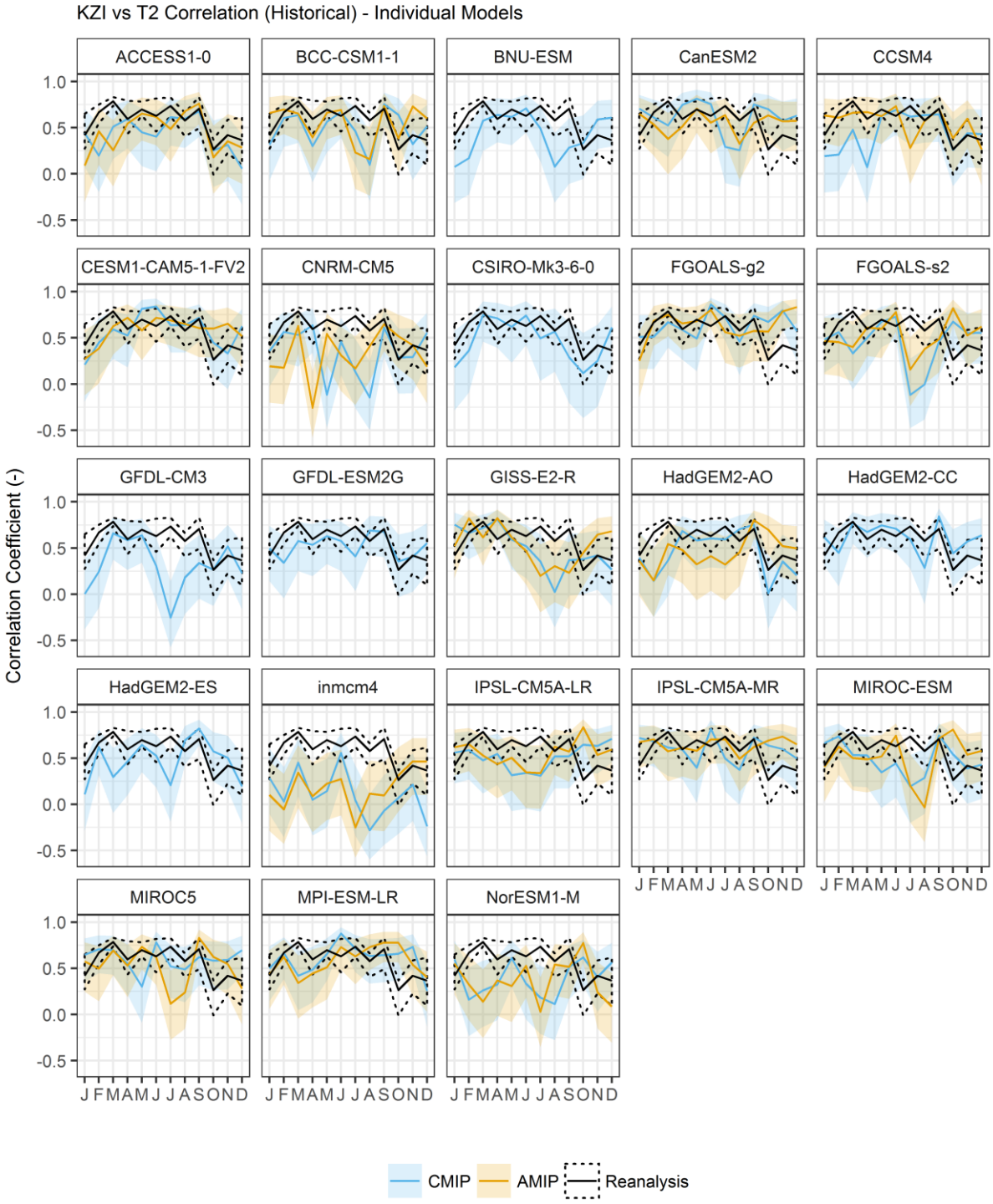


Figure C.1 – Comparison of KZI vs T2 correlations in CMIP and AMIP ensembles with reanalysis ensemble. Lines show means and shaded ranges (dashed lines) show 95% confidence intervals.

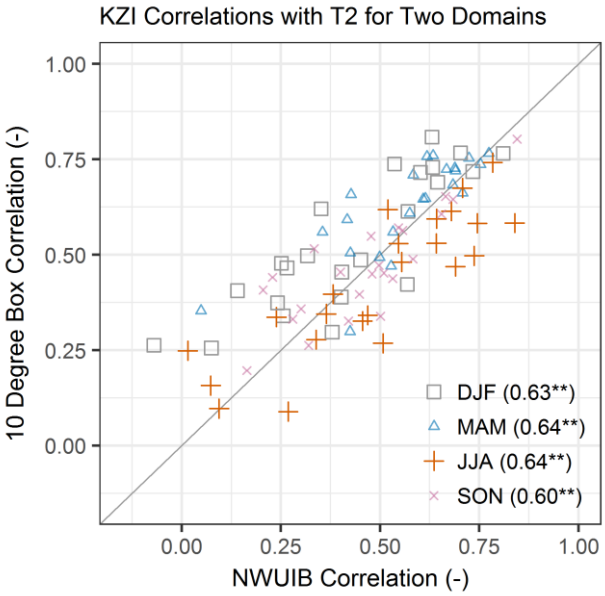


Figure C.2 – Comparison of KZI and temperature (2m) correlations for the NWUIB and 10° box domains. Numbers in brackets are Kendall's tau correlations, with ** representing significance at the 95% interval.

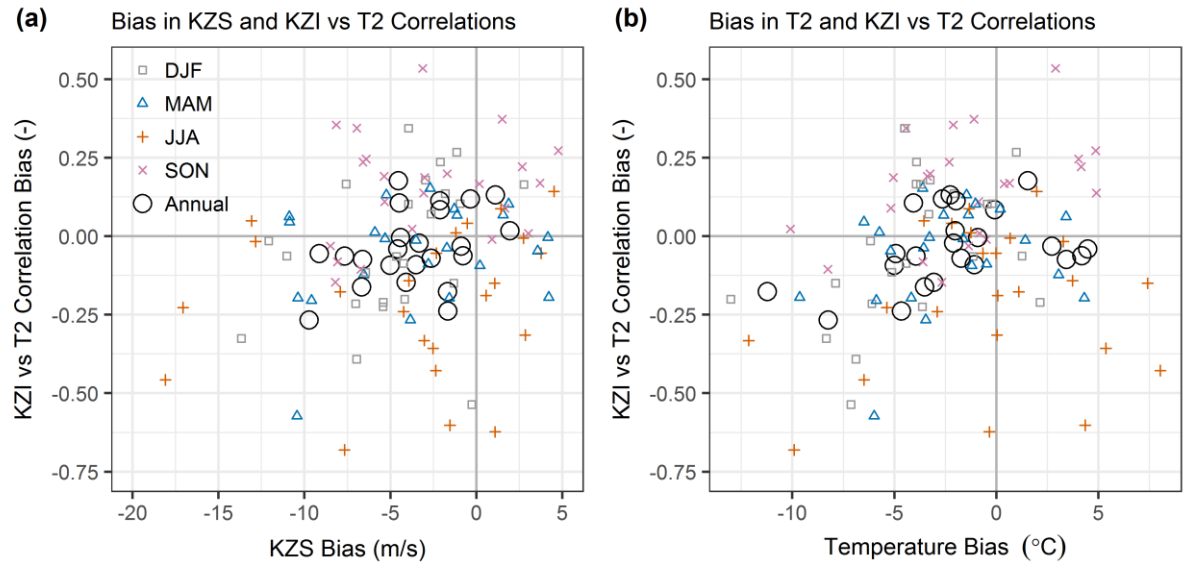


Figure C.3 – Comparison of KZS and KZI vs T2 correlation, both normalised by reanalysis ensemble means.

Season	KZS		Temperature (T2)	
	CMIP	AMIP	CMIP	AMIP
DJF	0.33**	0.04	0.36**	0.34*
MAM	0.12	0.19	0.15	0.12
JJA	0.19	0.18	0.03	0.15
SON	0.22	0.13	0.29*	0.16

Table C.1 – Correlation between biases in KZS, 2m temperature (T2) and KZI/T2 relationship (NWUIB).

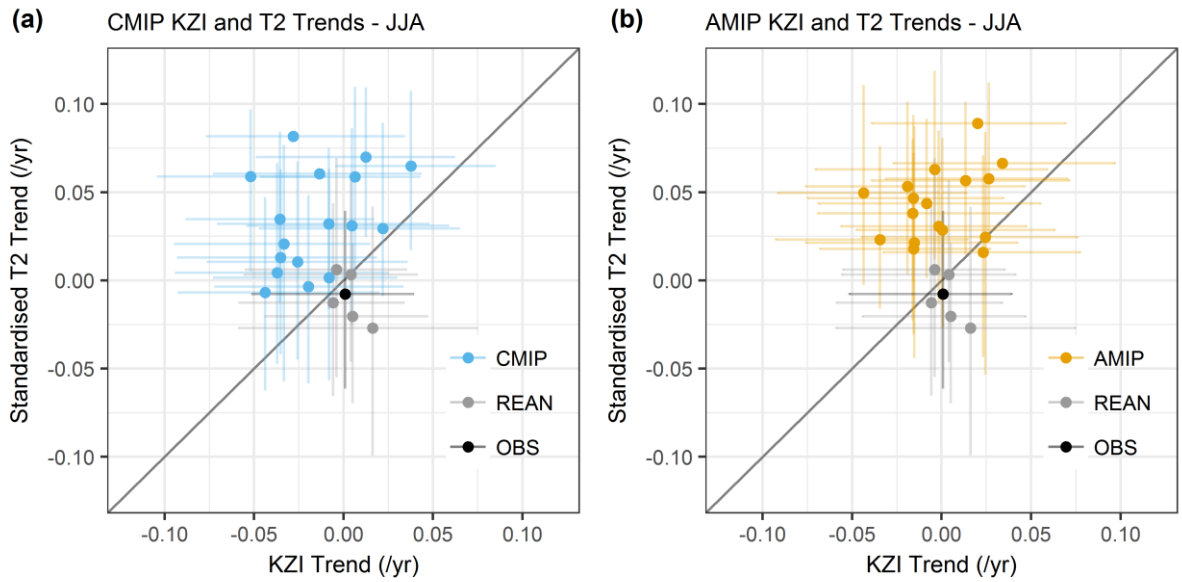


Figure C.4 – Comparison of (a) CMIP and (b) AMIP trends in KZI and T2 for the overlapping period of record (1979-2005).

Appendix D

Supporting Information for Chapter 7

D.1 Change Factors for the Distributions of Inter-Annual Variability

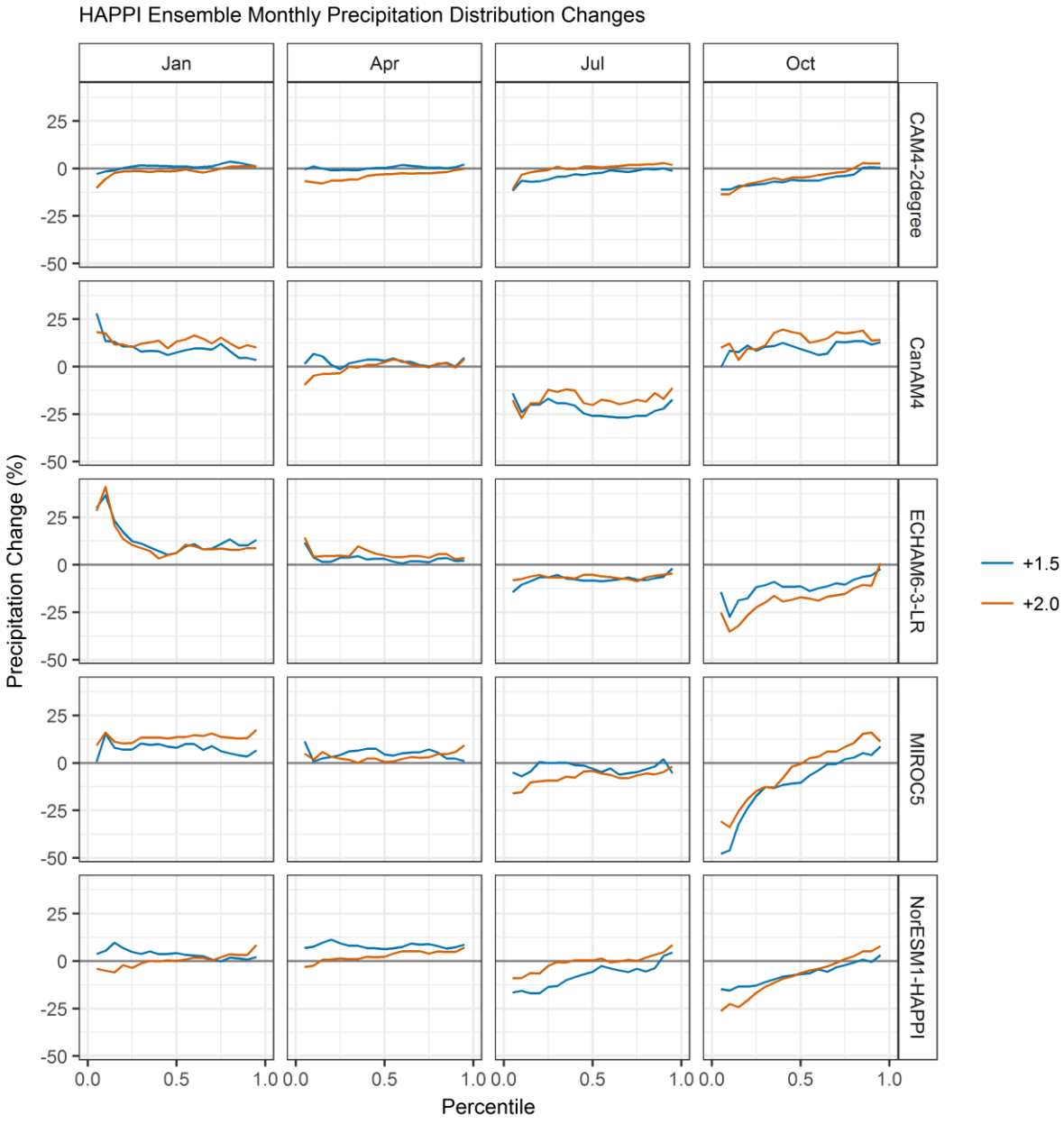


Figure D.1 – Precipitation change factors for different percentiles of the inter-annual distribution for specified months for the HAPPI models.

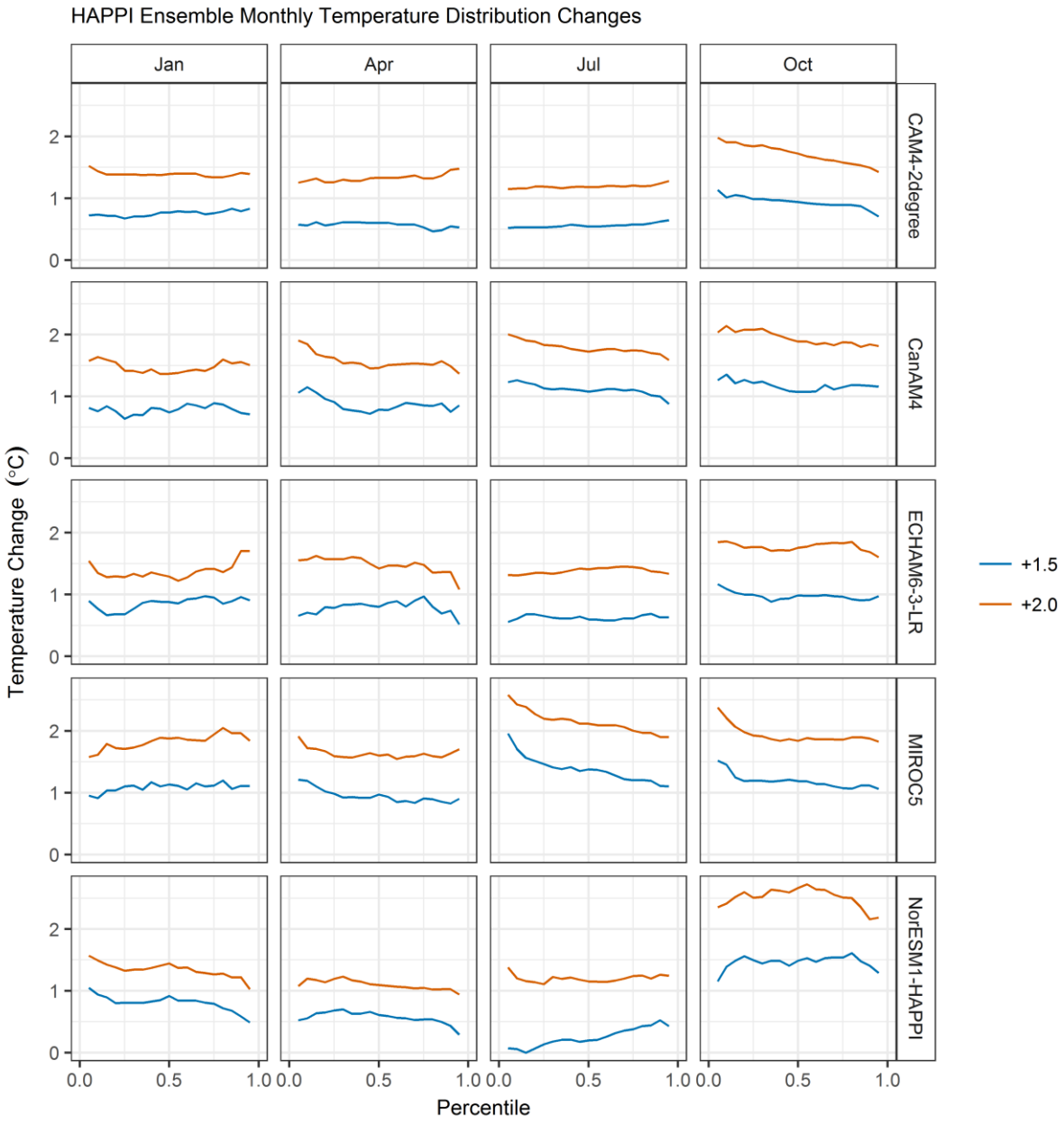


Figure D.2 – As Figure D.1 but for near-surface air temperature.

D.2 HAPPI Evaluation – High Resolution CAM5.1.2 Model

Temperature

Figure D.3a demonstrates that the annual cycle of monthly mean temperatures in CAM5.1.2 model exhibits a large amplitude when compared with both observations and the HAR. While only HAR10 is shown on Figure D.3a, the HAR10 and HAR30 annual temperature cycles are very similar (see Chapter 3). As such, and given the notable difference between CAM5.1.2 and HAR10, this suggests that the difference in annual cycles may not arise solely as a function of spatial resolution, because of the similarities between the HAR30 and CAM5.1.2

grid spacing. Figure D.3b shows that this large annual temperature cycle is also associated with a pronounced seasonal variation in bias relative to station observations. As it accounts for elevation differences between model cells and station locations, Figure D.3b suggests that there is a large cold bias in CAM5.1.2 in winter. The bias reduces during spring, such that by summer it approaches zero, although some variation between stations remains. This fundamental pattern is evident regardless of whether a constant (6.5°C/km) or model-derived, monthly lapse rate climatology is used to adjust the simulated time series to station elevations. The difference in bias as a result of methodological choice is thus likely to be small compared with the amplitude of the annual cycle of temperature bias.

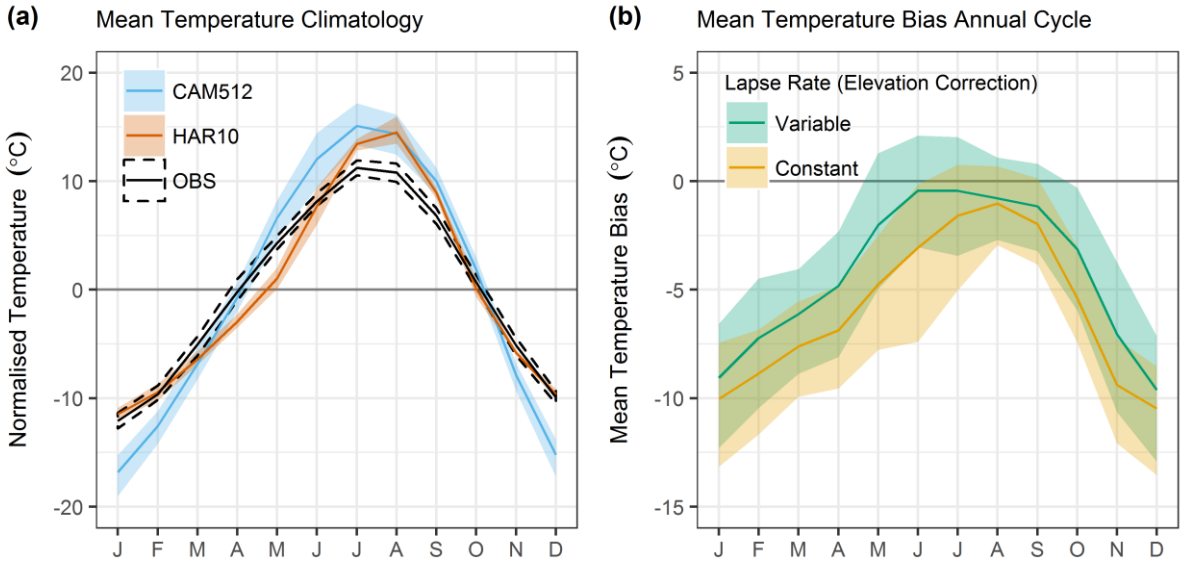


Figure D.3 – Evaluation of mean temperature climatology in CAM5.1.2. Annual cycles of monthly mean temperature are shown for CAM5.1.2, station observations and HAR10 in (a). Before summarising the station ensemble in (a), temperatures are normalised by subtracting the annual mean. The annual cycle of monthly mean temperature bias at station locations is in shown in (b). The two series show the implications of using time-varying lapse rates derived from CAM5.1.2 and a constant 6.5°C/km lapse rate to correct for elevation differences between model cells and station locations. Lines and shading in both panels show ensemble means and ranges, respectively

Figure D.4a compares the diurnal temperature range (DTR) simulated by CAM5.1.2 with observations and the HAR at station locations. Only HAR10 is shown, but Chapter 3 demonstrates that HAR10 and HAR30 have similar annual cycles of DTR. The most striking feature of Figure D.4a is that the DTR cycle for CAM5.1.2 is effectively the inverse of observations and the HAR. Unlike the HAR and observations, the largest DTR in CAM5.1.2 occurs in winter, rather than in summer. Figure D.4b suggests that this is primarily driven by

a more pronounced annual cycle of (monthly means of) daily minimum (night-time) temperatures compared with daily maximum (day-time) temperatures. This is apparent from the greater divergence of the minimum and maximum temperatures during winter months. We note that, even though the HAR representation of DTR is not in perfect agreement with observations, it does at least show an increase from winter to summer. This begs the questions of why DTR in CAM5.1.2 is so large and why minimum (night-time) temperatures diverge so much from maximum (day-time) temperatures in winter.

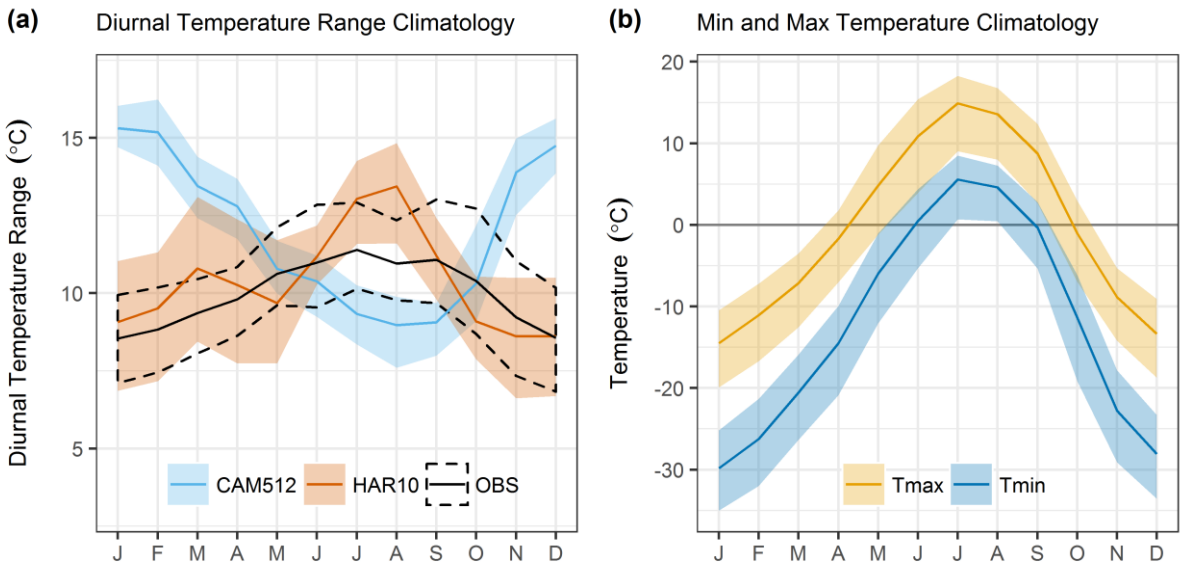


Figure D.4 – Evaluation of annual cycles of (a) diurnal temperature range and (b) daily minimum and maximum temperatures in CAM5.1.2. CAM5.1.2 is compared with HAR10 and observations in (a), while the monthly averages of daily minimum and maximum temperature at station locations in (b) are for CAM5.1.2 only. Lines and shading denoting the station ensemble mean and ranges, respectively

While Figure D.3 and Figure D.4 suggest notable seasonal variation in temperature bias in CAM5.1.2, Figure D.5 shows that mean temperature lapse rates in CAM5.1.2 tend to be notably steeper than those inferred from the HAR and observations. Again, Chapter 3 suggests that HAR10 and HAR30 show reasonably similar lapse rates, with both in more agreement with observations than HAR10. Moreover, the magnitude of intra-annual variation in CAM5.1.2 is comparatively large. The annual cycle shows two peaks, the first in May/June and the second in October/November, while the shallowest lapse rate is found in August. This seasonal variation is broadly present in the HAR and observations, even if their month-to-month variation and precise timing differs.

As lapse rates are at their steepest in spring and autumn, it may well be that a signal from the retreat and growth of snow-covered area (SCA) is strongly implicated in the lapse rate cycle in all datasets. If this is the case, it raises the question of why steeper lapse rates across the snowline occur in CAM5.1.2 compared with other model products, such as the HAR. Preliminary analysis (not shown) suggests that temperature biases are larger at higher elevations in the spring and autumn seasons in CAM5.1.2, but relatively constant across the elevation range in winter and, to a lesser degree, summer. This would again likely implicate snow cover as a possibly significant source of error, given that snow persists for longer at higher elevations in the spring and autumn seasons, while it extends across much of the UIB in winter. In conjunction with the fact that the temperature bias is much reduced in summer, this suggests that the representations of processes related to the surface energy balance and/or land-atmosphere fluxes over snow applied in CAM5.1.2 may need further investigation for applications in this region. However, initial analyses suggested that there might be errors in the aggregated model outputs for some of the energy balance components, so it is not clear whether new simulations would actually be required to investigate this fully.

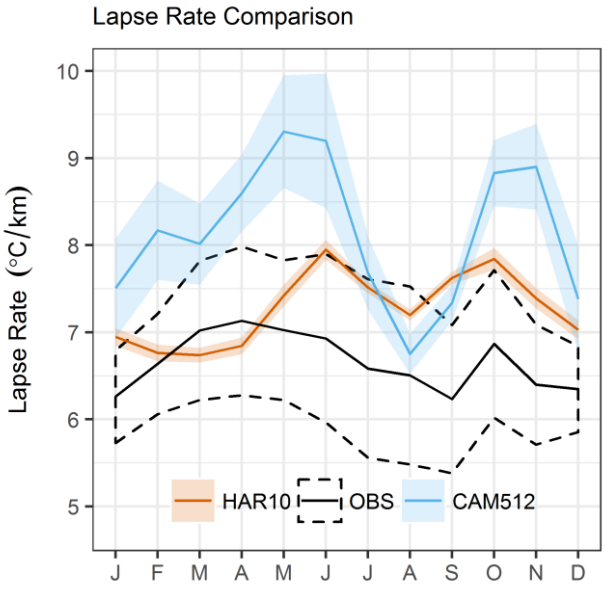


Figure D.5 – Comparison of annual cycle of monthly mean temperature lapse rates in CAM5.1.2, HAR10 and observations. Lapse rates were estimated using linear regression from station data for observations, while model outputs for cells within the NWUIB domain were used for CAM5.1.2 and HAR10. Ranges show 95% confidence intervals.

Returning to the issue of an ostensibly large DTR in winter, one possible contributor could be a deficiency in cloud cover. This might occur through excessive radiative cooling at night, if

cloudiness were underestimated. To investigate this, Figure D.6 compares the cloud cover fraction climatology from CAM5.1.2 with the MOD06L2 product. Interestingly, this suggests that the cloud cover representation in CAM5.1.2 may in fact be quite reasonable. The model's cloud cover fraction is a little lower compared with MODIS in some winter months, as well as in late spring and summer. Yet, overall the pattern seems reasonable, comparing favourably with the equivalent analysis of the HAR in Chapter 3. We note that this evaluation is based solely on mean cloud cover, without consideration of any diurnal asymmetry, as sub-daily outputs are not available for CAM5.1.2. However, the MODIS data products suggest day-time and night-time cloud cover annual cycles are really quite similar for the NWUIB domain (Forsythe et al. 2015). As such, if day-time and night-time annual cycles are also similar in CAM5.1.2, it could be that cloud cover is not at the root of the large winter DTR in the model. Delving further into this is beyond the scope of this work, but future efforts might again consider snow representation.

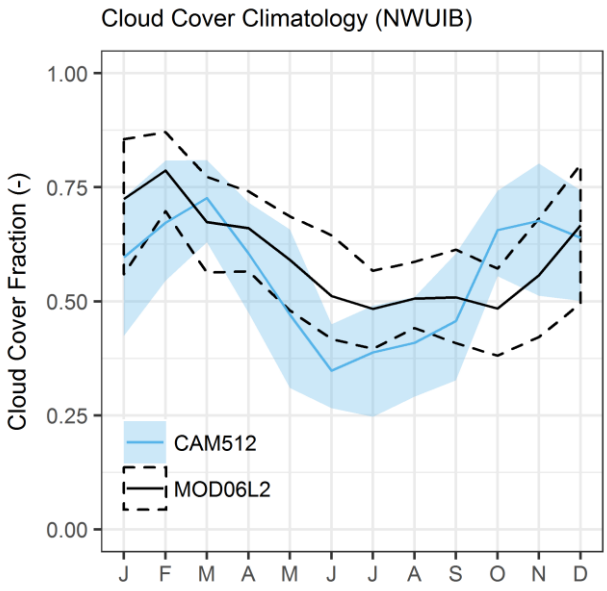


Figure D.6 – Comparison of CAM5.1.2 and MODIS (MOD06L2) mean cloud cover (NWUIB). Lines and ranges denote means and 10^{th}-90^{th} percentile ranges from inter-annual variation, respectively.

Precipitation

Figure D.7a describes the annual cycle of precipitation in the NWUIB for CAM5.1.2, station observations and the HAR. Firstly we note that the HAR agrees with observations insofar as it simulates a generally higher fraction of annual precipitation occurring in winter/spring compared with summer. As discussed in Chapter 3, the HAR appears to underestimate the

fraction of precipitation occurring in summer, as well as inter-station variability. In contrast, CAM5.1.2 exhibits a relatively diminished annual cycle in Figure D.7a, with comparatively moderate intra-annual variation. A smaller fraction of annual precipitation occurs in the December to February period in CAM5.1.2, but from March onwards fractional precipitation is reasonably constant on average, although inter-station variability grows in the summer months. This suggests that the model does not fully capture the relative balance between winter/spring snowfall brought by westerly disturbances and summer precipitation from monsoon offshoots or westerly-tracking depressions.

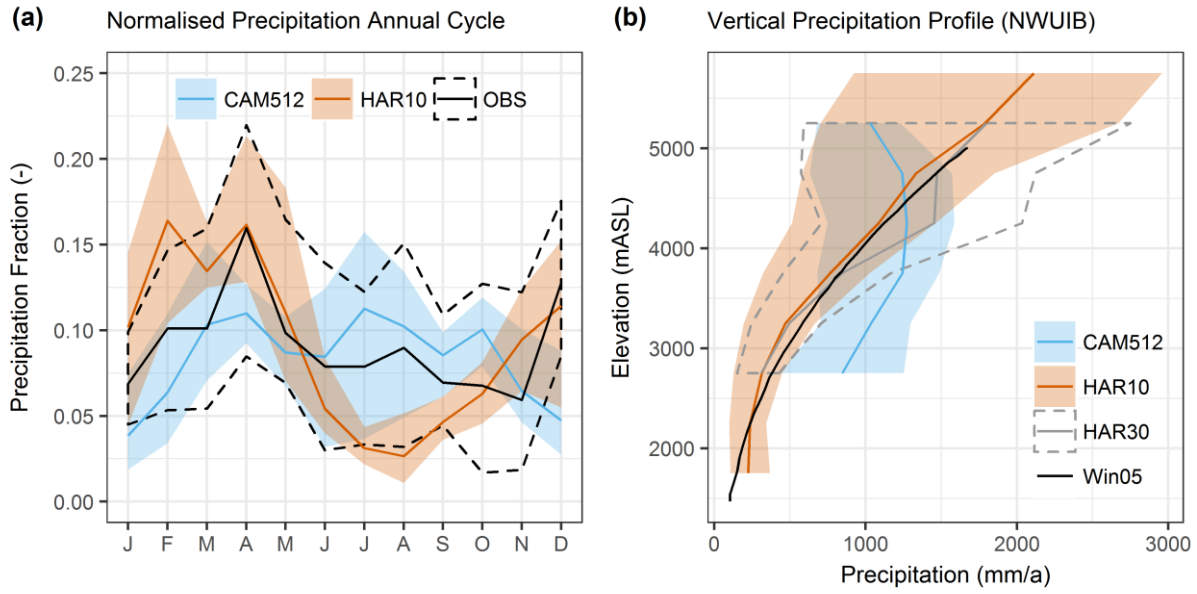


Figure D.7 – Evaluation of CAM5.1.2 precipitation climatology. Annual cycles of precipitation expressed as the fraction of the annual total falling each month at the station locations are given in (a) for CAM5.1.2, HAR10 and observations. Lines and shading show the mean and range from the station ensemble. Elevation profiles for mean annual precipitation are given in (b) for CAM5.1.2, HAR10, HAR30 and the observation-derived profile (Win05) of Winiger et al. (Winiger et al. 2005). Shading denotes inter-quartile range of model cells within the NWUIB domain for CAM5.1.2 and the HAR.

For an absolute comparison, Figure D.7b shows vertical profiles of precipitation for the NWUIB for CAM5.1.2, the two HAR products, and an observational-based curve (Winiger et al. 2005). This indicates some differences in form between CAM5.1.2 and the HAR. CAM5.1.2 exhibits comparatively high precipitation at lower elevations (less than ~4000 mASL) compared with the HAR, which matches relatively well with Winiger et al.'s profile, especially in the case of HAR10. At elevations above ~4000 mASL, CAM5.1.2 precipitation begins to decrease overall, while HAR precipitation continues to increase with elevation. While precipitation is often considered to decrease above a critical elevation in the UIB (see

Chapter 3), this is typically thought to be around 5000-5500 mASL, rather than the ~4000 mASL level in CAM5.1.2. It is also notable that the HAR30 profile agrees more closely with HAR10 than the CAM5.1.2 profile. As HAR30 is at a comparable resolution to CAM5.1.2, this suggests that the discrepancies in the latter's elevation profiles may not be a simple function of resolution differences relative to the higher resolution HAR10 product. Possible avenues to consider may be the implications of transferring "tuning" from coarser to high resolution in the model. Indeed, Wehner et al. (2014) demonstrate that performance in some respects is degraded in the 0.25° version of CAM5 relative to its 1° counterpart.

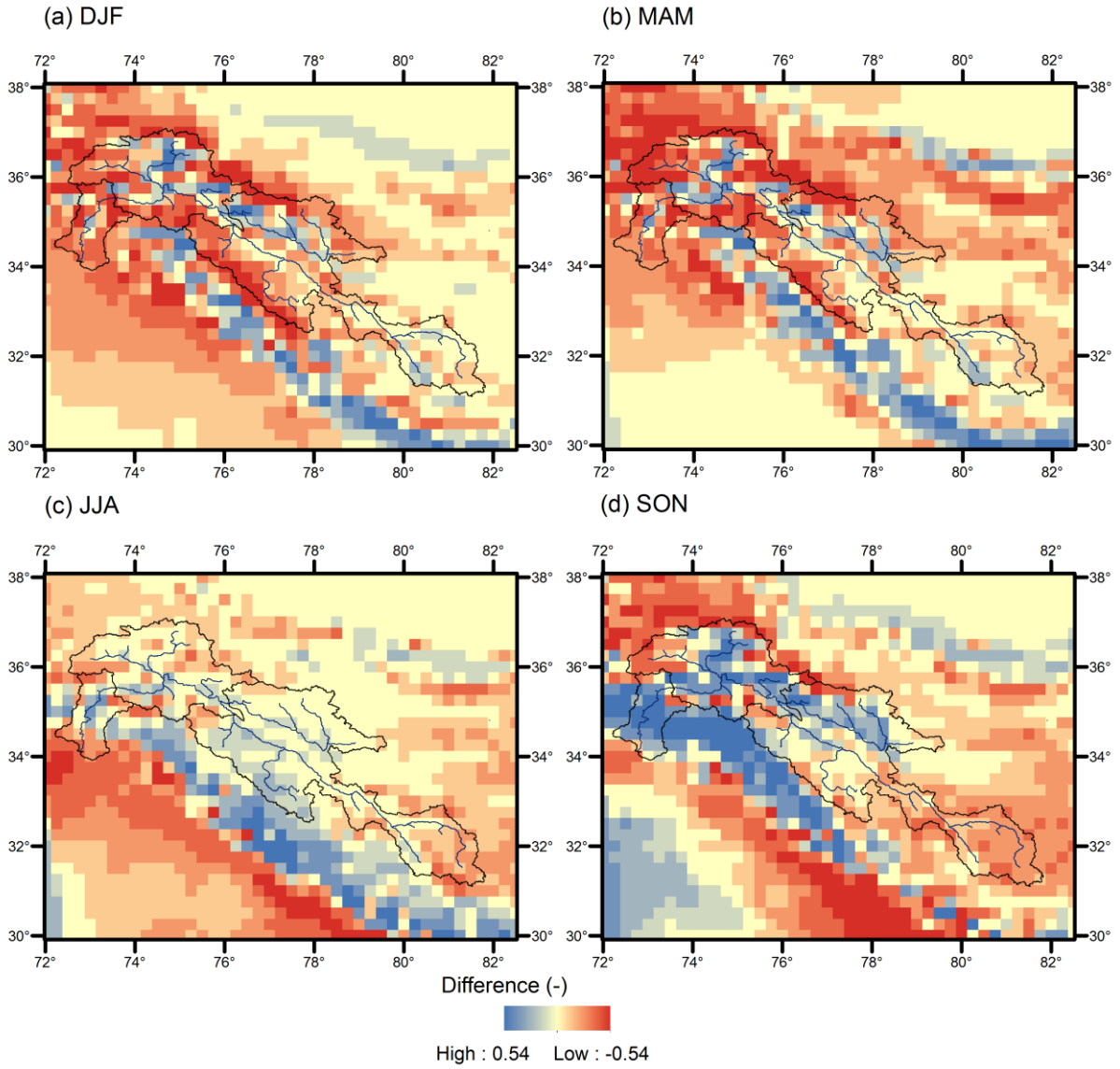


Figure D.8 – Difference in normalised seasonal mean precipitation between CAM5.1.2 and HAR10, with the latter re-gridded to match the 0.25° resolution of CAM5.1.2. Normalisation was carried out separately for each dataset by scaling the modelled values to the 0-1 based on the (spatial) minimum and maximum values in the window surrounding the UIB. Differences were then calculated as CAM5.1.2 minus HAR10.

To provide a slightly broader view of CAM5.1.2's precipitation fields, Figure D.8 shows the spatial distribution of differences in normalised seasonal precipitation between CAM5.1.2 and HAR10, after the latter was re-gridded to the CAM5.1.2 resolution. Normalisation was undertaken by scaling the precipitation in each model cell to the range 0-1 using the minimum and maximum values from all cells in the window surrounding the UIB shown in Figure D.8. This helps to indicate the underlying similarity of the spatial distributions of precipitation in each season by removing some of the influence of biases in magnitude.

In each season, Figure D.8 suggests that there is a tendency for CAM5.1.2 to simulate relatively higher precipitation along the "front ranges" situated upwind of the UIB with respect to the prevailing westerlies. The front ranges follow the linear band of positive differences (blue in Figure D.8) oriented north-west to south-east, which sits to the south-east of the UIB. In winter (DJF) and spring (MAM), there is also a tendency for CAM5.1.2 to over-predict the relative precipitation in major valleys in the NWUIB, as well in the major Shyok tributary. This accords with the over-estimation at low elevations in Figure D.7b, as does the underestimation (in relative terms) of precipitation at higher elevations in the NWUIB. In summer (JJA), the differences between CAM5.1.2 and HAR10 within the NWUIB and UIB actually tend to reduce somewhat, potentially reflecting a reduction in the strength, coherence and magnitude of orographic gradients in summer precipitation conditions in the basin. Yet, by autumn (SON), there is major disagreement along the front ranges forming the south-east side of the UIB. There is also widespread over-estimation of relative precipitation within the NWUIB and Shyok domains again.

Of course the HAR is itself a model dataset and so an imperfect reference, but it is interesting to note the differences in the spatial distribution of precipitation between the two data products. The strong relationship between orography and divergences between the models implied by Figure D.8 suggests that the processes underpinning orographic precipitation could need further refinement in CAM5.1.2 if it were to provide WRF-like behaviour. Indeed, this may be desirable, given that WRF shows great promise for simulating precipitation in the data-sparse Himalayan region (see e.g. Chapter 3 and Maussion et al. (2014)). In conjunction with the temperature analysis, this evaluation of the HAPPI CAM5.1.2 model therefore suggests that, despite its high resolution, it exhibits a number of deficiencies with respect to simulating UIB climatology. As such, and given the smaller ensembles available for the model relative to the other HAPPI models, we suggest that further analysis and potentially refinement of CAM5.1.2 would be useful before application in this region. It is not clear that

its high resolution adds much at this point, such that we omit it from the remainder of the chapter.

D.3 CMIP5 Change Factors by Model

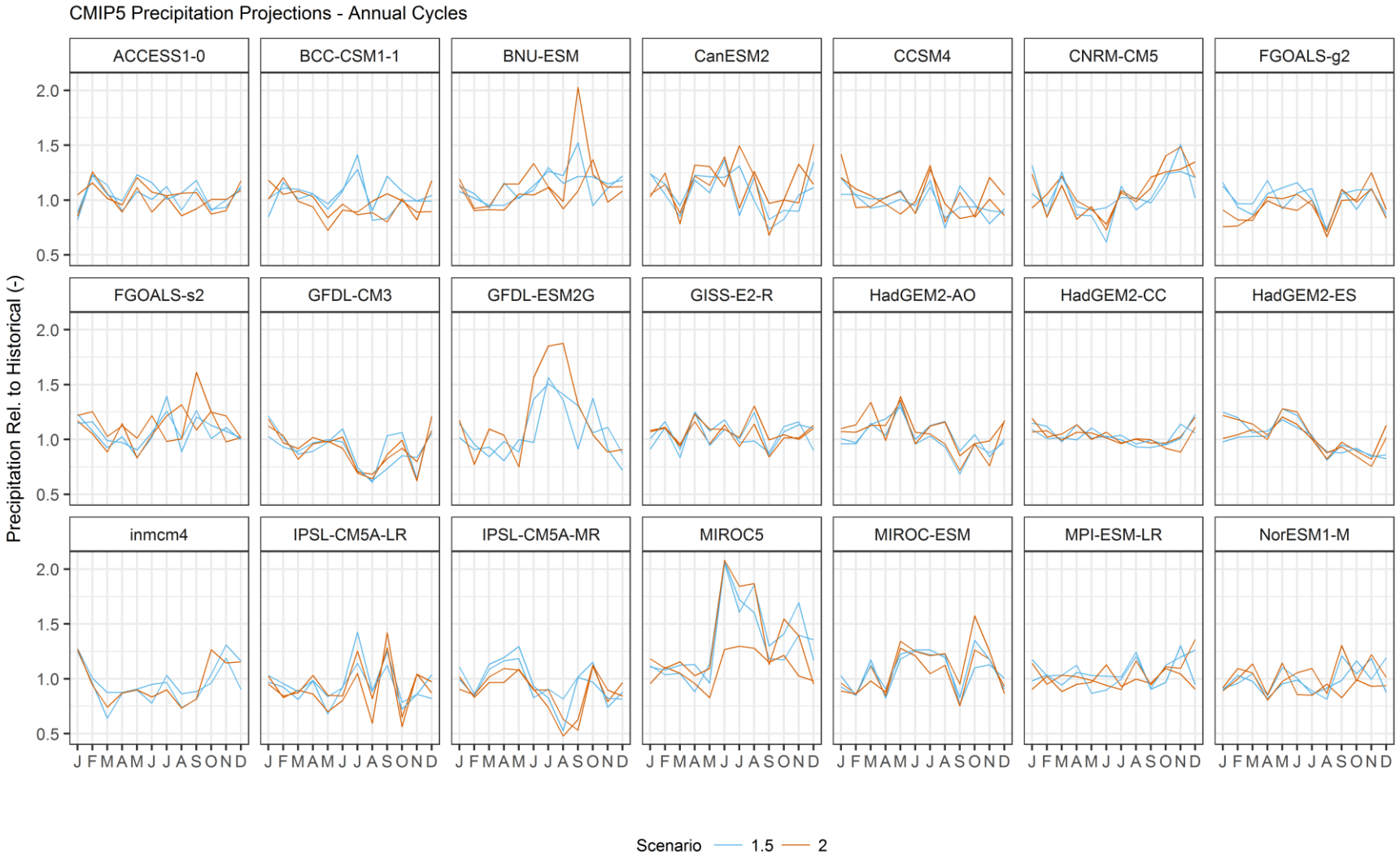


Figure D.9 – Change factors for the monthly precipitation climatology in CMIP5 models

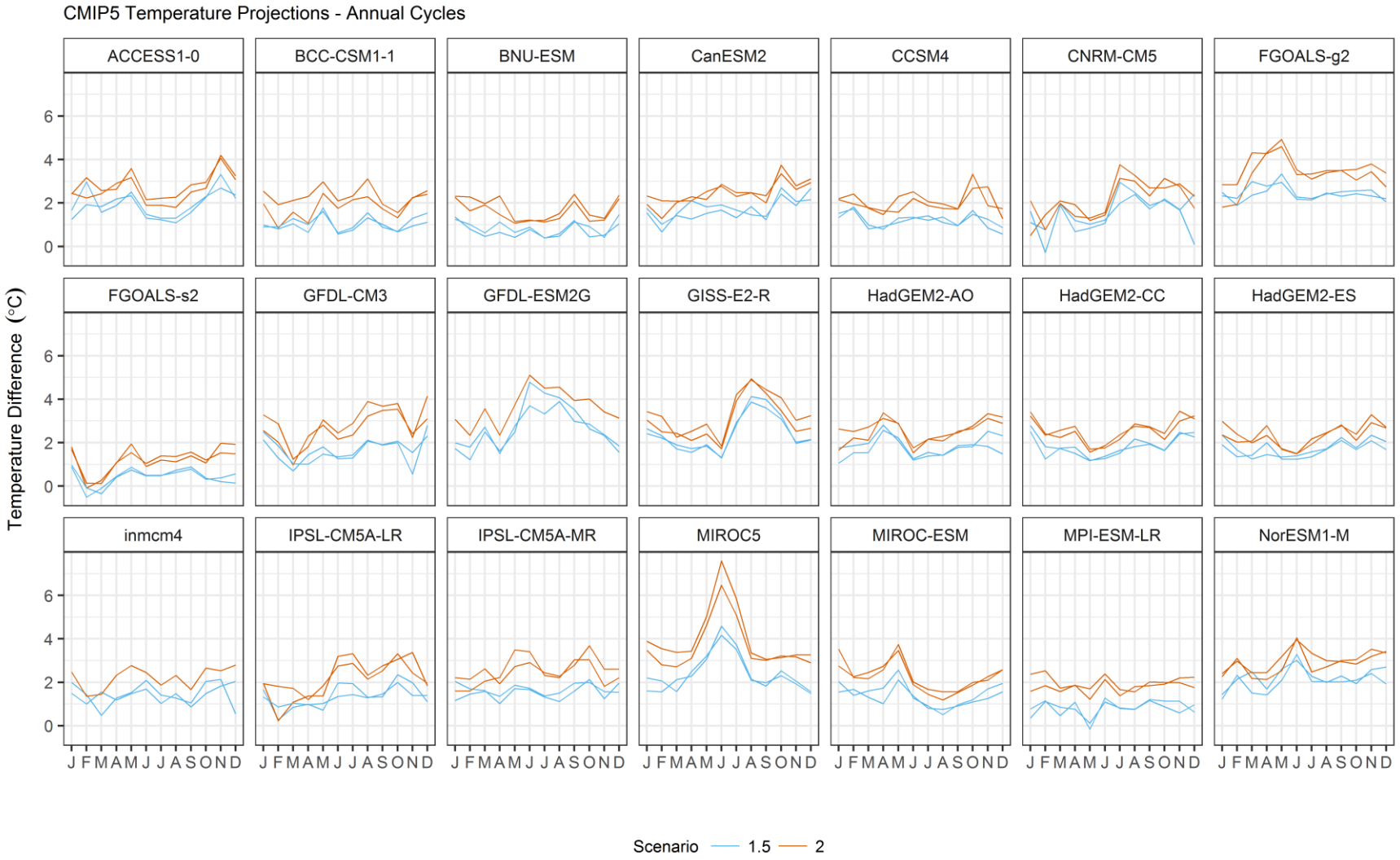


Figure D.10 – As Figure D.9 but for near-surface air temperature.

D.4 Higher Temperature Scenarios in CMIP5

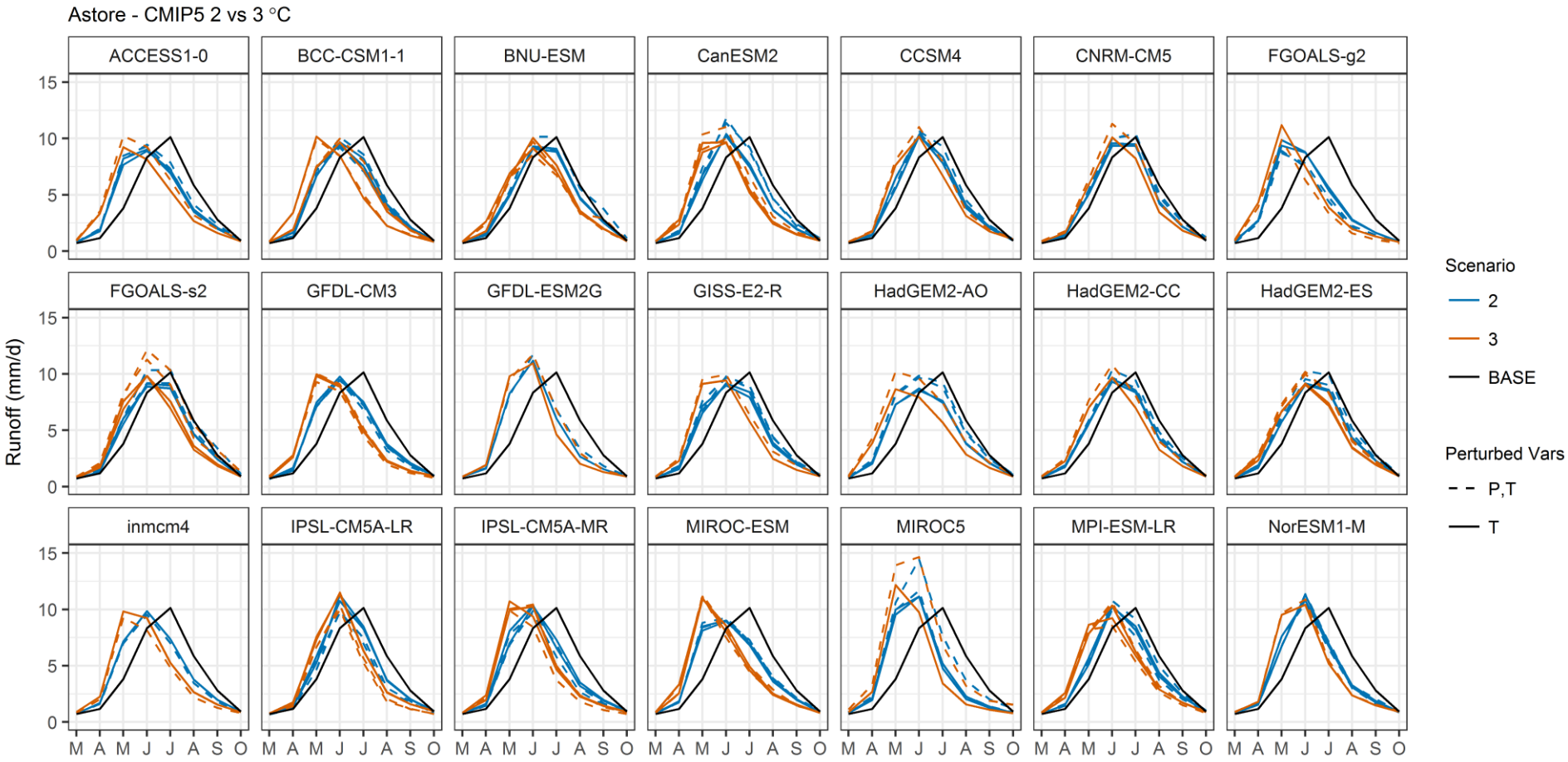


Figure D.11 – Simulated mean monthly flows in TOPKAPI-ETH for 2 and 3°C warmer worlds for the Astore sub-basin, based on change factors from CMIP5. Runs using change factors calculated separately for RCP4.5 and RCP8.5 are both shown. The responses are also differentiated by runs including temperature-only perturbations and runs considering both precipitation and temperature perturbations.

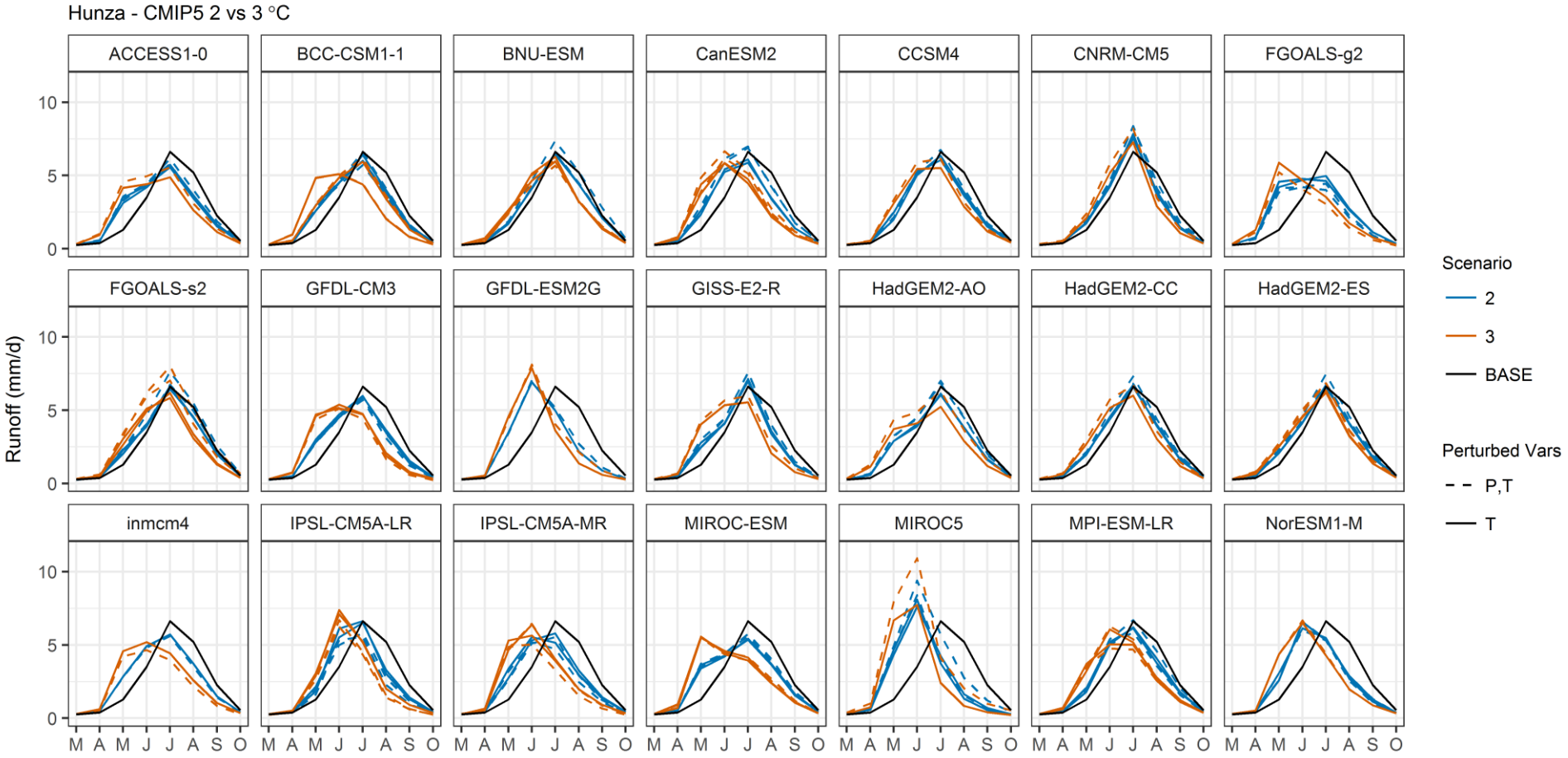


Figure D.12 – As Figure D.11 but for the Hunza sub-basin.