Electronic Thesis and Dissertation Repository

5-12-2016 12:00 AM

A Physical Scaling Method to Incorporate Regional Physical Characteristics in Future Climatic and Hydrologic Predictions

Abhishek Gaur
The University of Western Ontario

Supervisor
Dr. Slobodan P. Simonovic
The University of Western Ontario

ABSTRACT

In this thesis an original Physical Scaling (SP) method for downscaling Global Circulation Model (GCM) based climatic projections has been developed, tested and applied over a study region. The model formulation can take into account regional physical characteristics like land-cover and elevation into the model formulation. A thorough verification of the method and its extension: SP with Surrounding pixel information (SPS) method has been performed and their performance towards downscaling GCM based precipitation, surface temperature and air temperature has been compared with many state-of-the-art downscaling models like Bias Correction Spatial Downscaling (BCSD) method, Statistical DownScaling Method (SDSM) and Generalized Linear Modeling (GLM). The SPS method extends SP method by also taking into account neighborhood physical characteristics into the downscaling process. A major benefit of the presented downscaling approaches is that they can account for non-stationarity in physical characteristics of the region of interest like changes in land-cover as well as their neighborhoods. This represents a major contribution in the field of statistical downscaling literature since it brings the benefits of physically based dynamic downscaling into a statistical downscaling framework.

Proposed models are used to isolate physically sourced climatic and hydrologic contributions in four catchments located within the southern Saskatchewan region of Canada. Contributions towards flood magnitudes are also studied for low to high return period flooding events. Results indicate that the contributions of catchment physical characteristics towards shaping climatic and hydrologic regimes in the analyzed catchments are statistically significant. Further significant variability in the detected changes exists over catchment space and analyzed time-period.

Finally the results from this thesis highlight the importance of further exploration of physically driven climatic changes, and the need to find out how to incorporate them while making future streamflow predictions. The developed SP and SPS methods are highly relevant and useful in a non-stationary world which is set to experience rapid climatic and geophysical changes in the future.

Keywords: Downscaling, SP method, SPS method, Statistical downscaling, Physical scaling, Climate change.

CO-AUTHORSHIP

This thesis is prepared in accordance with the regulations for Integrated-Article (Formerly Manuscript) format thesis stipulated by the School of Graduate and Postdoctoral Studies at The Western University. Work towards all the papers listed below is performed by the author under the supervision of Dr. Slobodan P. Simonovic. The co-authorship of chapters 2-5 is provided below:

Chapter 2: A Scaling Method for Physically Representative Downscaling of Climate Model Data.

Authors: Abhishek Gaur and Slobodan P Simonovic

Submitted for publication in the *Climate Dynamics* journal.

Contributions

Author AG designed the study, performed the statistical analysis, wrote the protocol, and wrote the first draft of the manuscript. Author SPS assisted the study design, supervised the analyses, reviewed the first draft of the manuscript and suggested revisions. Both authors read and approved the final manuscript.

Chapter 3: Extension of SP method and its application towards downscaling climate model based near surface air temperature.

Authors: Abhishek Gaur and Slobodan P Simonovic

Submitted for publication in the *Journal of Applied Meteorology and Climatology*.

Contributions

Author AG designed the study, performed the statistical analysis, wrote the protocol, and wrote the first draft of the manuscript. Author SPS assisted the study design, supervised the analyses, reviewed the first draft of the manuscript and suggested revisions. Both authors read and approved the final manuscript.

Chapter 4: Application of Physical Scaling towards downscaling climate model precipitation data.

Authors: Abhishek Gaur and Slobodan P Simonovic

Submitted for publication in the *Theoretical and Applied Climatology* journal.

<u>Contributions</u>

Author AG designed the study, performed the statistical analysis, wrote the protocol, and wrote the first draft of the manuscript. Author SPS assisted the study design, supervised the analyses, reviewed the first draft of the manuscript and suggested revisions. Both authors read and approved the final manuscript.

Chapter 5: Physically sourced climatic changes and their implications on future flow projections.

Authors: Abhishek Gaur and Slobodan P Simonovic

Submitted for publication in the *Hydrological Processes* journal.

<u>Contributions</u>

Author AG designed the study, performed the statistical analysis, wrote the protocol, and wrote the first draft of the manuscript. Author SPS assisted the study design, supervised the analyses,

reviewed the first draft of the manuscript and suggested revisions. Both authors read and approved the final manuscript.

To my Teachers, Parents and Bhagwan

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

List of acronyms

AT	air temperature
AD	aqua day
AN	aqua night
AR5	intergovernmental panel for climate change fifth assessment report
AV	atmospheric variables
BCBI	bias-correction and bilinear interpolation
BCSD	bias correction and spatial disaggregation
BI	bilinear interpolation
BSV	barren or sparsely vegetated
C	croplands
CDF	cumulative distribution function
CF	change factor
CMIP5	coupled model intercomparison project – phase 5
COV	coefficient of variation
CS	closed shrublands
D.mod	downscaling model
DNF	deciduous needleleaf forest
DBF	deciduous broadleaf forest
ddays	fraction of dry days
E	elevation
ENF	evergreen needleleaf forest
EBF	evergreen broadleaf forest

fnf	forest or no forest
G	grasslands
GAM	generalized additive model
GEV	generalized extreme value
GLM	generalized linear modelling
GLM.sig	generalized linear modelling with initial selection of predictors
GWR	geographically weighted regression
hcdc	high cloud area fraction
hgt	geopotential height
IAM	integrated assessment model
IPCC	intergovernmental panel for climate change
lc	land-cover
lcdc	low cloud area fraction
LHZ	land-use harmonization
LR	linear regression
LST	land surface temperature
masl	meters above sea level
MODIS	moderate-resolution imaging spectroradiometer
MEBWG	multisite, multivariate, maximum entropy bootstrap weather generator
MF	mixed forest
mcdc	medium cloud area fraction
NARR	north american regional reanalysis
NASA	national aeronautics and space administration

NCEP	national center for environmental protection
OS	open shrublands
p1d	total number of one-day precipitation events
p2to4d	total number of 2-4 day precipitation events
p5d	total number of 5 or more day precipitation events
PCM	NCAR-DOE parallel climate model
ppt.max	maximum precipitation intensity
ppt.wet	mean wet day precipitation
PRD	pearl river delta
PSC	physically sourced changes
QR	quantile regression
RDAS	regional data assimilation system
RMSE	root mean squared error
RP	return period of flood event
S	savannas
sc	snow covered
SDSM	statistical downscaling model
SDSM.sig	statistical downscaling model with initial selection of predictors
sf	snow free
shum	specific humidity
sma	secondary mean age
sp.cor	spearman correlation coefficient
SP	physical scaling

SP-LC	SP method without land-cover as predictor
SP-LC-elev	SP method without elevation and land-cover as predictors
SPS3x3	SPS method with 3x3 neighborhood scale
SPS5x5	SPS method with 5x5 neighborhood scale
SPS7x7	SPS method with 7x7 neighborhood scale
SPS9x9	SPS method with 9x9 neighborhood scale
SR	spatial robustness
T	temperature
TD	terra day
TERC-RAMS	terrestrial environment research center - regional atmospheric modeling system
TN	terra night
TR	temporal robustness
TRMM	tropical rainfall measuring mission
UB	urban and built-up
UHI	urban heat island
UMD	university of maryland
UMD-LHZ	university of maryland-land use harmonization
URI	urban heat island ratio
W	water
WRF	weather research and forecasting
WS	woody savannas
WSC	water survey of canada

CHAPTER 1: Introduction

In this chapter the theoretical background of the problem that is addressed in this thesis is provided. First a brief introduction to the phenomenon of climate change is provided and the role of Global Circulation Models (GCMs) in the context of climate change is discussed. Next a brief introduction to downscaling methods is provided. A discussion on the relative merits and demerits of statistical and downscaling methods is provided. This is followed by a highlight of the historically observed land-cover and elevation effects on climate. The motivation for the development of SP and SPS downscaling methods is described thereafter. The section ends with a description of the layout of the thesis and a list of theoretical contributions made in this research.

1.1 Global climate change and GCMs

It is now well established and scientifically documented that climate change has been observed historically and that anthropogenic greenhouse gas emissions are a major source contributing to this phenomenon (Bindoff et al. 2013). Significant changes in climatic and hydrologic variables have been recorded across the globe in general (Hartmann et al. 2013) and Canada in particular (refer to results of for instance Gaur and Simonovic 2015; Das and Simonovic 2013; Grillakis et al. 2012). Changes in climatic and hydrologic extremes occurring due to climate change have also been well documented (Seneviratne et al. 2012).

GCMs and Earth System Models (ESMs) are mathematical representations of the earth's climate system and can simulate complex bio-geophysical and bio-chemical cycles that shape global climate (Taylor et al. 2012; Flato 2011). They are used as tools to study the interaction of greenhouse gases with earth's climate system as well as to simulate and predict future

climate response under future greenhouse gas scenarios. In the Intergovernmental Panel for Climate Change (IPCC) Fifth Assessment Report (AR5) a total of 39 GCMs and ESMs have been identified (Flato et al. 2013). Several climate model inter-comparison projects like Coupled Model Intercomparison Project-Phase 5 (CMIP5) have been initiated to facilitate the intercomparison and evaluation of GCMs (Taylor et al. 2012). Climate models have been found to evolve continuously with time in terms of their accuracy and reliability (Sun et al. 2015; Grose et al. 2014; Yao et al. 2013).

1.2 Downscaling methods

Owing to the limitations in computational resources, typical spatial scale at which GCM outputs are generated is typically more than 1° x 1° (or approximately 110 Km x 110 Km). This spatial scale is very large for regional and catchment scale impact assessment studies. The process of estimating higher resolution data from low resolution GCM data is termed downscaling. Two broad classes of downscaling methods have been identified: statistical and dynamical. In statistical downscaling methods, higher resolution climate data are estimated by developing statistical relationships between large scale climatic or atmospheric data and locally observed data and those relationships are used to predict future local climate. In dynamic downscaling methods local scale climate is estimated by coupling a mesoscale higher resolution model or Regional Climate Model (RCM) with the GCM. Boundary conditions generated from the GCMs are used as inputs into the RCM to estimate local climate in a physically based way. In this thesis the term physically based is associated with climate projections obtained from models that simulate physical processes occurring within the earth's biosphere.

Both statistical and dynamic downscaling methods have been extensively used in the past to perform climate change impact assessment studies (Xue et al. 2014; Schoof 2013). Both methods have certain advantages and disadvantages. For instance some of the widely known issues with statistical downscaling approaches are: a) they are not physically based, b) their application is inherently dependent on the assumption that the statistical relationship between large scale climate processes and local scale climate will remain unchanged in future thereby limiting their reliability to climates that are similar to what are being observed currently c) they assume similar geophysical structure of the study region between calibration and prediction time-periods; and d) their application is problematic in data scarce regions. There are however many advantages of using them as well. Some of the advantages of using statistical methods are: a) they are computationally inexpensive and produce reasonable results very fast so uncertainty associated with future emission scenarios as well as climate models can be addressed using them; b) they are very important in cases where topography or sub-grid scale features (<1 km) play a significant role in shaping the local climate. c) Results from statistical methods can be benchmarked to specific sites and site-specific changes can be analyzed. On the other hand, all physically based models are gridded and hence sub-grid scale uncertainty is not addressed in their projections, d) real-time operations can be performed using them since they can be quickly executed. Dynamical downscaling methods, on the other hand, are physically based but require significant computational resources. Therefore their use in real-time operation or uncertainty assessment is very limited.

1.3 Local climate variability with land-cover and elevation

According to the Lowry's model (Lowry 1977) local scale climate depends on large scale processes, local scale land-cover and elevation properties. A change in any of the above mentioned factors will impact the local scale climate. Although Lowry identified these variables for shaping landscape scale climate, many studies have found significant variations of local climate with changes in land-cover at a city scale. For instance after analyzing Urban Heat Island (UHI) trends between 1989 and 2010, Qiao et al. (2014) concluded a 200% increase in the Urban Heat Island Ratio (URI) for the Beijing city. They found that the URI for the city, which quantifies contribution rate of urban land towards UHI development, has not only intensified but also has expanded spatially with increasing urban sprawl on suburban areas. Similar results were obtained by Hu and Jia (2010) after analyzing changes in UHI magnitude in greater Guangzhou (China) from 1980-2007. They found that the mean Land Surface Temperature (LST) increased by 3.1 K in the city during this duration. UHI magnitude increased in intensity as well as spatial extent as the surrounding cropland areas were subjected to urban sprawl. These studies show that the climatic behavior of a location is influenced by its land-cover. The science and physical processes responsible for this behavior has been well documented in UHI literature (Voogt and Oke 2003; Oke 1982; Oke 1987). A very comprehensive review of observed and modelled evidences of the relationship between land-cover and climate is also provided in Pielke et al. (2011).

Several observational studies have noted statistically significant influences of land-cover on precipitation at a city scale. For instance Li et al. (2011) studied urban signature in strong and weak precipitation events over the Pearl River Delta (PRD) in China using Tropical Rainfall Measuring Mission (TRMM) satellite precipitation data. They found that over and around the

urban regions "strong" precipitation events have increased with urbanization while "weak" precipitation events have decreased. They also found strengthening of the precipitation intensity, a decrease in rainfall frequency, an increase in convective rainfall and afternoon precipitation events over and around the urban areas. Similar findings were reported by De and Rao (2004). They analyzed rainfall trends of several Indian megacities (with population more than 1 million) such as: New Delhi, Kolkata, Mumbai and Chennai between 1901 and 2000 and found statistically significant increasing trends in annual and monsoon precipitations. A decreasing trend was also found for a few cities. They found more pronounced increases in precipitation during the period 1951-2000 when rapid industrial development took place over the selected urban locations. Rao et al. (2004) performed a similar analysis on precipitation trends for the duration 1901-2000 and found similar statistically significant increasing trends for the cities analyzed. Kishtawal et al. (2010) analyzed mean and extreme rainfall trends of urban locations within India using observed as well as remotely sensed TRMM precipitation data and identified an increasing trend linked to the pace of urbanization of the cities. Further urban locations were found to have more possibility of witnessing an extreme precipitation event than the surrounding non-urban area. Several other studies have also found evidences of land-cover linkage with rainfall (Kug and Ahn 2013; Schluzen et al. 2010; Halfon et al. 2009; Fujibe et al. 2009; Mote et al. 2007; Diem and Mote 2005; Inoue and Kimura 2004; Dixon and Mote 2003; Shepherd et al. 2002). There are three hypotheses as to how urban areas can impact regional precipitation distribution: a) by modifying the thermodynamic processes such as energy balance and urban heat island induced circulation within and around the city, b) by causing winds to converge over and downwind of the cities due to roughness of the city elements, and c) by effecting

cloud microphysical processes due to the presence of large amounts of aerosol in the urban air (Han et al. 2014).

Relationship between elevation and rate of temperature increase with greenhouse gas concentrations has been evaluated in some studies. It has been found that the rate of temperature change is higher at high elevations than at the lower elevation regions. For instance Yan and Liu (2014) found that higher elevation areas (> 2000 masl) in the Tibetan Plateau show a higher rate of warming than other lower elevation regions surrounding it. However there are other studies which have not found any elevation related association of warming rates. A detailed review of the observational evidence and plausible operating mechanisms that lead to these elevation dependent responses to greenhouse gases have been detailed in Pepin et al. (2015). Climate modification brought due to snow-albedo feedback, more frequent cloud cover, and water-vapor related radiative feedbacks are considered as possible mechanisms for a higher warming rate in the higher elevation regions. The relationship of precipitation with elevation has also been studied. For instance Puvaneswaran and Smithson (1991) found both increasing and decreasing trends while analyzing precipitation-elevation relationships across Sri Lanka and termed the relationship to be complex. Using Geographically Weighted Regression (GWR), Brunsdon et al. (2001) found a definite relationship between elevation and precipitation over the Great Britain. They highlighted the importance of considering GWR while studying these relationships as they vary in space. Lastly it has been recognized that the conclusions made for elevation dependent changes in climate variables are uncertain because of less data availability at higher elevation regions (Pepin et al. 2015; Rangawala and Miller 2012).

1.4 Motivation for the development of Physical Scaling method

One of the major shortcomings of statistical downscaling methods as discussed in section 1.2 is that they are not physically based. Therefore in this research attempt is made to include important physical parameters into a statistical downscaling framework so that the downscaled outputs are physically representative if not physically based. In this thesis, the term physically representative is associated with downscaled projections that are able to simulate the variations of climate with physical factors like elevation, land-cover and their distribution across the region of interest. In section 1.3 it is highlighted that local scale climate is influenced by the land-cover and elevation related physical properties of the study region. Therefore these two physically based variables are used as explanatory variables in the SP method formulation. Both elevation and land-cover are considered in the model definition to account for climatic changes caused by changes in the combination of both these factors. Further in SPS method land-cover and elevation configuration of the surrounding pixels are also taken into consideration.

Annual 500 m land-cover data for the period 2002-2012 provided by the MODerate-resolution Imaging Spectroradiometer (MODIS) and 90 m elevation product from National Aeronautics and Space Administration (NASA) Shuttle Radar Topographic Mission (SRTM) provided necessary data for model calibration and validation.

1.5 Research objectives and theoretical contributions

1.5.1 Research objective

The objective of this research is to develop and implement a generic methodology for the identification of future physically sourced climatic and hydrologic changes. The

term physically sourced changes signify changes that are caused by changes in physical characteristics of a region like changes in land-cover and elevation distribution of a region. The research attempts to answer following research questions:

- Traditional statistical downscaling models are purely statistical in nature. Is it possible to make statistical downscaling procedure physically representative by including physical parameters like elevation and land-cover into the model definition?

- How does such a physically representative statistical downscaling model perform in the period where observed precipitation and temperature records are available?

- In Hurtt et al. (2011) coarsely gridded land-use estimates are provided for the period 1500-2100 at annual time-steps. Is it possible to derive high resolution future land-cover projections from these land-use projections?

- Can high resolution future climate projections be inferred by making use of developed high resolution land-cover maps and physically representative statistical downscaling model?

- What is the contribution of physically sourced climatic and hydrologic changes towards shaping future climatic and hydrologic conditions? What are the factors that significantly influence these changes?

1.5.2 Theoretical contributions

- This research has added a new dimension to statistical downscaling process. Statistical downscaling models have previously been developed for a location and hence could only account for changes in large scale climate. The proposed Physical

Scaling model is unique in the sense that it is first statistical downscaling model that includes physical parameters in its definition thereby providing an opportunity to address non-stationarity in both large scale climate system as well as local scale physical characteristics of a region while making future climatic projections.

- This research outlines and implements a methodology (in Chapter 5) which can be adopted to derive physically driven climatic and hydrologic response within a catchment. The developed methodology is generic and can be adopted to estimate physically driven climatic and hydrologic projections at any catchment located across the globe.

- A methodology to downscale and reconfigure gridded land-use projections into land-cover projections has been described in Chapter 2 of this thesis. This is a generic methodology which can be used to develop future land-cover maps at any region of the world using the land-use projections from Hurtt et al. (2011).

- The indirect approach of downscaling air temperatures as outlined in Chapter 3 can be very helpful in performing downscaling of GCM based air temperature projections in data sparse regions. Typically statistical downscaling methods are location specific and hence can only be used at locations where they are calibrated. Since SP and SPS methods are calibrated using physical parameters as predictors, the relationship can is transferable over space and time as evident from the results in Chapter 3 where these models are found to perform well in both spatial and temporal robustness tests.

- This study also improves upon the existing methods that were used to predict air temperatures from surface temperatures. It is found in Chapter 3 that the inclusion

of atmospheric variables in this model improves the model efficiency by over 35% towards predicting air temperatures from surface temperatures.

1.6 Layout of the thesis

This thesis is divided into six chapters. The current chapter is followed by a set of four chapters:

Chapter 2: The SP method is introduced in this chapter. Further its performance towards downscaling reanalysis based gridded surface temperature data is evaluated and model performance is compared with the performance of a state-of-the-art downscaling model: BCSD. The model is thereafter used to obtain future surface temperature projections in the study region.

Chapter 3: SP method is extended in this chapter and SPS method is introduced. Also two different approaches: direct and indirect, towards downscaling air temperature data are detailed. An ensemble of SP method based models with two different methods (SP and SPS), different approaches (direct and indirect) and different functional forms (linear regression, quantile regression and Generalized Additive Models) are evaluated for their ability to downscale reanalysis based near surface air temperature. Further the impact of the choice of different methods, approaches and functional forms on future temperature projections is quantified and compared.

Chapter 4: In this chapter, SP and SPS method based models are evaluated for their ability to downscale reanalysis based gridded precipitation data. The model performance is compared with two other state-of-the-art downscaling models: GLM approach and SDSM. The

validated models are thereafter used to make future precipitation projections in the study region.

Chapter 5: In this chapter, SP and SPS method based models are used to study precipitation, temperature and outflow at four catchments located across the southern Saskatchewan region of Canada. The aim is to ascertain if physically sourced climatic and hydrologic changes are statistically significant or not. Further it was of interest to find factors that influence these changes. The impact of physically sourced climatic changes on flood magnitudes is also quantified in this chapter.

Chapter 6: In this chapter primary conclusions made from the results obtained in chapters 2-5 are summarized and future direction of work is identified.

References

Bindoff NL, Stott PA, AchutaRao KM, Allen MR, Gillett N, Gutzler D, Hansingo K, Hegerl G, Hu Y, Jain S, Mokhov II, Overland J, Perlwitz J, Sebbari R, Zhang X (2013) Detection and Attribution of Climate Change: from Global to Regional. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker TF, Qin D, Plattner GK, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

De US, Rao GSP (2004) Urban climate trends - The Indian scenario. J. Ind. Geophys. Union 8: 199-203.

Diem JE, Mote TL (2005) Interepochal changes in summer precipitation in the southeastern United States: Evidence of possible urban effects near Atlanta, Georgia. J. Appl. Meteorol. 44: 717-730.

Dixon PG, Mote TL (2003) Patterns and causes of Atlanta's urban heat island-initiated precipitation. J. Appl. Meteorol., 42: 1273-1284.

Flato GM (2011) Earth system models: an overview. WIREs Clim Change 2:783–800.

Fujibe F, Togawa H, Sakata M (2009) Long-term change and spatial anomaly of warm season afternoon precipitation in Tokyo. Sci. Online Lett. Atmos. 5: 17-20.

Grose MR, Brown JN, Narsey S, Brown JR, Murphy BF, Langlais C, Gupta AS, Moise AF, Irving DB (2014) Assessment of the CMIP5 global climate model simulations of the western tropical Pacific climate system and comparison to CMIP3. Int. J. Climatol. 34: 3382–3399.

Halfon N, Levin Z, Alpert P (2009) Temporal rainfall fluctuations in Israel and their possible link to urban and air pollution effects. Environ. Res. Lett. 4: 025001, doi:10.1088/1748-9326/4/2/025001.

Han J-Y, Baik J-J, Lee H (2014) Urban impacts on precipitation. Asia-Pacific Journal of Atmospheric Sciences 50(1): 17-30.

Hartmann DL, Klein Tank AMG, Rusticucci M, Alexander LV, Brönnimann S, Charabi Y, Dentener FJ, Dlugokencky EJ, Easterling DR, Kaplan A, Soden BJ, Thorne PW, Wild M, Zhai PM (2013) Observations: Atmosphere and Surface. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker TF, Qin D, Plattner GK, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Hu Y, Jia G (2010) Influence of land use change on urban heat island derived from multi-sensor data, International Journal of Climatology 30: 1382-1395.

Inoue T, Kimura F (2004) Urban effects on low-level clouds around the Tokyo metropolitan area on clear summer days. Geophys. Res. Lett. 31: L05103, doi:10.1029/2003GL018908.

Kishtawal CM, Niyogi D, Tewari M, Pielke RA Sr., Shepherd JM (2010) Urbanization signature in the observed heavy rainfall climatology over India. Int. J. Climatol. 30: 1908-1916.

Kug J-S, Ahn MS (2013) Impact of urbanization on recent temperature and precipitation trends in the Korean peninsula. Asia-Pac. J. Atmos. Sci. 49: 151-159.

Li W, Chen S, Chen G, Sha W, Luo C, Feng Y, Wen Z, Wang B (2011) Urbanization signatures in strong versus weak precipitation over the Pearl River Delta metropolitan regions of China. Environ. Res. Lett. 6: 034020, doi:10.1088/1748-9326/6/3/034020.

Lowry WP (1977) Empirical estimation of the urban effects on climate: A problem analysis. J. Appl. Meteor. 16:129–135.

Mote TL, Lacke MC, Shepherd JM (2007) Radar signatures of the urban effect on precipitation distribution: A case study for Atlanta, Georgia. Geophys. Res. Lett. 34: L20710, doi:10.1029/2007GL031903.

Oke, T. R., 1982: The energetic basis of the urban heat island. Quart. J. Roy. Meteor. Soc., 108, 1–24.

Oke, T. R., 1987: Boundary Layer Climates. 2nd ed. Routledge, 435 pp.

Pepin N, Bradley RS, Dias HF, et al. (2015) Elevation-dependent warming in mountain regions of the world. Nature Climate Change 5: 424-430.

Pielke RA, Pitman A, Niyogi D, et al. (2011) Land use/land cover changes and climate: modeling analysis and observational evidence. WIREs Clim Change 2:828–850.

Qiao Z, Guangjin T, Lixiao Z, Xu X (2014) Influences of Urban Expansion on Urban Heat Island in Beijing during 1989–2010. Advances in Meteorology 187169, 11 pages.

Rao GSP, Jaswal AK, Kumar MS (2004) Effects of urbanization on meteorological parameters. MAUSAM 55: 429-440.

Schlünzen KH, Hoffmann P,Rosenhagen G, Riecke W (2010) Long-term changes and regional differences in temperature and precipitation in the metropolitan area of Hamburg. Int. J. Climatol. 30: 1121-1136.

Seneviratne SI, Nicholls N, Easterling D, Goodess CM, Kanae S, Kossin J, Luo Y, Marengo J, McInnes K, Rahimi M, Reichstein M, Sorteberg A, Vera C, Zhang X (2012) Changes in climate extremes and their impacts on the natural physical environment. In: Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation [Field CB, Barros V, Stocker TF, Qin D, Dokken DJ, Ebi KL, Mastrandrea MD, Mach KJ, Plattner GK, Allen SK, Tignor M, Midgley PM (eds.)]. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change (IPCC). Cambridge University Press, Cambridge, UK, and New York, NY, USA, pp. 109-230.

Shepherd JM, Pierce H, Negri AJ (2002) Rainfall modification by major urban areas: Observations from spaceborne rain radar on the TRMM satellite. J. Appl. Meteorol. 41: 689-701.

Sun Q, Miao C, Duan Q (2015) Comparative analysis of CMIP3 and CMIP5 global climate models for simulating the daily mean, maximum, and minimum temperatures and daily precipitation over China, J. Geophys. Res. Atmos.120: 4806–4824.

Taylor KE, Stouffer RJ, Meehl GA (2012) An overview of CMIP5 and the experimental design. Bulletin of the American Meteorological Society 93: 485–498.

Voogt JA, Oke TR (2003) Thermal remote sensing of urban areas. Remote Sensing of Environment (special issue on Urban Areas) 86: 370-384.

Yan L, Liu X (2014) Has climatic warming over the Tibetan Plateau paused or continued in recent years? J. Earth Ocean Atmos. Sci. 1: 13–28.

CHAPTER 2: A Scaling Method for Physically Representative Downscaling of Climate Model Data

2.1 Introduction

Global warming is expected to play a significant role in shaping future climatic conditions (Stocker et al. 2013). Further it is expected that local and regional scale physiography will respond to the changed global atmospheric forcing and produce complex climatic changes in future. General Circulation Models (GCMs) can simulate complex biophysical interactions occurring within the earth's climate system and are used to simulate climatic response to future greenhouse gas emissions. However, outputs generated by the GCMs are of low spatial resolution than that required for regional and local scale climate change impact assessment studies. The process of estimating local scale climate variable of interest from GCM simulated climatology is referred to as downscaling in the climate change impact assessment literature. Downscaling methodologies used in the past can be classified into two broad categories: a) statistical and b) dynamic methods, while some studies combine the two approaches (Fowler et al. 2007). Statistical downscaling methods employ statistical methods to link GCM simulated climatology and locally observed climate data. On the other hand, dynamic downscaling methods use a higher resolution physically based mesoscale model called Regional Climate Model (RCM) nested within a GCM to model climate at high spatial resolutions.

Apart from large scale atmospheric processes, land-cover and elevation are two important factors that are known to shape temperature patterns at local and regional scales (Lowry 1977; Oke 1982; 1987; Stewart 2000). Observational evidences of the influences of changing

land-cover on regional temperature patterns have been found. For instance after analysing historical Urban Heat Island (UHI) trends, Qiao et al. (2014) concluded a 200% increase in the Urban Heat Island Ratio (URI) of Beijing city between 1989 and 2010. They found that the URI for the city, which quantifies contribution rate of urban land towards UHI development, has not only intensified but also has expanded spatially with increasing urban sprawl on suburban areas. Similar results were obtained by Hu and Jia (2010) in the greater Guangzhou region in China. They found that between 1980 and 2007, mean Land Surface Temperature (LST) of the region increased by 3.1 K. UHI magnitude increased in intensity as well as spatial extent as the surrounding cropland areas were subjected to urban sprawl during this period. The results from these and other studies (Fall et al. 2010; Hale et al. 2006; 2008; Roth and Chow 2012; Ezber et al. 2006; Lemonsu et al. 2015) suggest that regional land-cover distribution significantly influences the local and regional temperature patterns and that changes in land-cover distribution affects regional and local climatology.

Projected future changes in land-cover should therefore be considered while making regional and local scale temperature projections. Several dynamic downscaling based studies have incorporated land-cover changes while making future temperature projections. For instance Argueso et al. (2014) used the Weather Research and Forecasting (WRF) modeling system to downscale CSIRO MK3.5 GCM outputs to 2 Km grid-scale. They simulated the present (1990-2009) and future (2040-2059) climates for the Sydney area and concluded that coupling of future urbanisation effects and climate change will significantly affect the local climatology of the city in future. They projected more intense increases in minimum temperatures than in maximum temperatures, particularly in winter and spring season. Adachi et al. (2012) also calculated future UHI intensities for Tokyo city by using five future

projections from climate models downscaled using the Terrestrial Environment Research Center - Regional Atmospheric Modeling System (TERC-RAMS) regional model. After comparing the results obtained with and without incorporating urban effects they concluded that the temperature change between 1990s and 2070s owing to greenhouse gas emissions is projected to be ~2 ℃ while that due to land-cover changes is ~0.5 ℃. Several other dynamical downscaling based studies (Georgescu et al. 2013; Kusaka et al. 2012; Hamdi et al. 2013; McCarthy et al. 2012) have also concluded similar changes in regional climatology in future.

The effects of land-cover and elevation have been largely ignored in the statistical downscaling literature. Statistical relationships derived while performing statistical downscaling are generally location specific. For instance, Salathe (2003) calculated precipitation scaling factors which were essentially ratios of the observed and National Center for Environmental Protection (NCEP) based modelled precipitation values, at each 50 Km x 50 Km grid-point located within the Yakima River basin, USA. These ratios were thereafter used with future GCM projections made by three climate models to obtain downscaled future precipitation across the study region in Salathe (2005). Wood et al. (2004) used Bias Correction and Spatial Disaggregation (BCSD) approach to first calculate a spatial anomaly pattern across all 1/8° x 1/8° grids located within the Columbia river basin, USA using observed and NCAR-DOE Parallel Climate Model (PCM). This spatial anomaly pattern was thereafter used to downscale future temperature and precipitation projections across the grids located within the study region. Gaur and Simonovic (2015) downscaled future precipitation and temperature at 52 gauging stations located within the Grand river basin (Ontario, Canada) using a weather generator approach. Change factors were calculated

at each gauging station using historical and future climate model data. These change factors were thereafter used to obtain downscaled precipitation and temperature data at each gauging station using Multisite, Multivariate, Maximum Entropy Bootstrap Weather Generator (MEBWG) (Srivastav and Simonovic 2014). It is assumed in these studies that the relationship between GCM data and local climate is a function of its location, and this relationship stays constant over time. However, as discussed before, land-cover and other physical parameters influence local and regional temperature patterns. Therefore with changes in these physical characteristics of a location, the relationship between locally observed and model based temperatures should also change.

To model such geophysical changes within a statistical downscaling framework, a Physical Scaling (SP) method is proposed in this study. This method is based on a hypothesis that local scale temperatures can be defined using large scale climate and land-cover, elevation properties of a location. If this hypothesis is true, the relationships developed can be used along with future projected climatic and land-cover projections to estimate local scale future temperatures. In this study, the above mentioned hypothesis is tested. The developed model is thereafter used to downscale surface temperatures across the southern Saskatchewan region of Canada. The rest of the paper is organized as follows. First a description of the study area is provided in section 2.2 followed by the datasets used in section 2.3. The methodology used to perform the analysis is described in section 2.4 followed by a description of the models used in this study in section 2.5. A discussion on the results obtained is provided in section 2.6. Lastly conclusions made from this study are summarized in section 2.7.

2.2 Study Area

The southern Saskatchewan region of Canada is selected as a case study area in this study. The political and physiographic settings of the region are shown in Figure 2.1. The study area is land-locked and encompasses many small lakes and streams. Altitude within the study region ranges from 240 masl to 1389 masl. Two major urban centers of Saskatoon and Regina are located within the study area. The climate of Saskatchewan is continental and is characterised by large fluctuations in temperature (up to 65ºC). These fluctuations are due to the land-locked position of the study region within the North American land-mass, because of which the region heats up as well as cools down quickly. An important climatic feature of the region is frequent clear skies and sunny conditions. The majority of precipitation that Saskatchewan receives occurs during summers due to the passing of mid-latitude cyclones over the region. Wintertime precipitation occurs as snow and due to sustained below zero temperatures accumulated snow-pack has a major influence on the climatology of the region (Encyclopedia of Saskatchewan 2015).
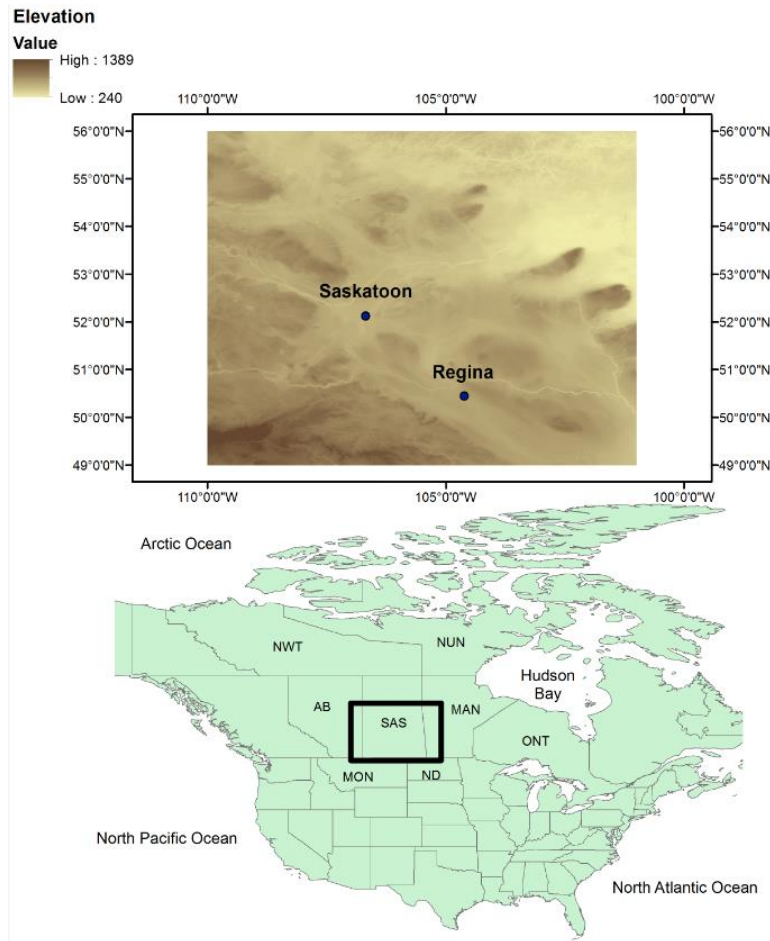
Figure 2.1. The political and physiographic settings of the area under study.

The study area is characterized by a multitude of different land-cover regions. Northern regions of the study area are dominated by the forested land-cover while the central and southern regions are dominated by the croplands and grassland land-cover. Overall cropland occupies the largest fraction (close to 50%) of the total area followed by the forested land (close to 30%). The annual remotely sensed land-cover distribution of the study region over the period 2006-2012 is shown in Figure 2.2. It can be seen that the land-cover composition of the region has not changed significantly over the period 2006-2012. Most significant changes are observed for land-cover classes: evergreen needle-leaf forest (+4%), croplands

(+3%), mixed forest (+1%), woody savannas (-3%) and grasslands (-4%). Average surface temperature across different land-cover classes are presented in Figure A1. It can be noticed that the temperature varies significantly across different land-cover classes.
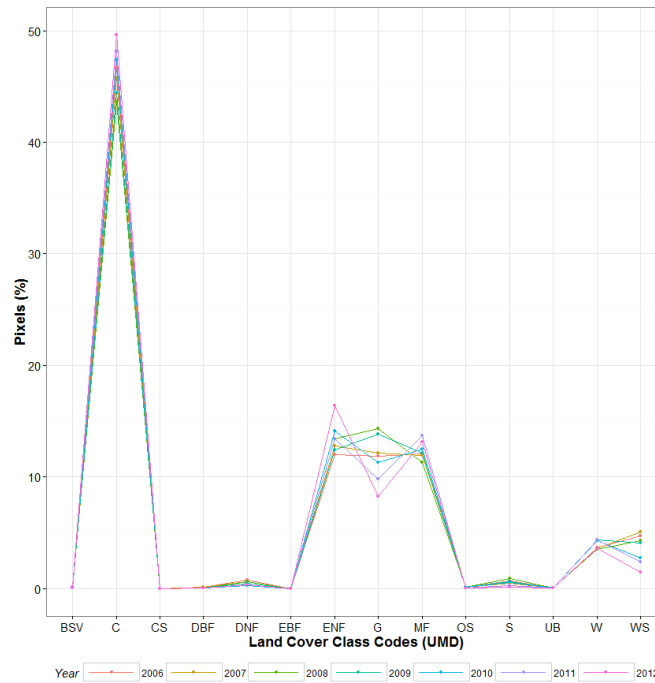
Figure 2.2. Annual land-cover distribution across the study region over the period 2006-2012.

2.3 Data

2.3.1 Moderate Resolution Imaging Spectroradiometer (MODIS) recorded surface temperature level 3 Terra (MOD11A1) and Aqua (MYD11A1) product. Terra passes equator at around 10:30 AM/PM while Aqua passes at around 1:30 PM/AM. Day-time as well as night-time surface temperature data products from both satellites available between 2006 and 2013 have been used. The data are available at approximately 1 Km x 1 Km spatial resolution. The total number of dates for which data has been analyzed

is equal to 9,383. The percentages of the total data available in each month and year of the study period are shown in Table 2.1. It can be seen that a higher percentage of reliable pixel data is available for the snow-free months (April to October) than the snow-covered months (November to March). The reason behind this can be that more cloud-free conditions occur during the summers than in winters facilitating the sensing of reliable surface temperature values by the satellites. The distribution of data is also uneven over the time of the day with higher percentage of data available in nights (67%) than days (33%).

Table 2.1 Distribution (%) of the remotely sensed surface temperature data over the period 2006-2013.

Year	Month											
	J	*F*	*M*	*A*	*M*	*J*	*J*	*A*	*S*	*O*	*N*	*D*
2006	0.8	0.7	0.8	1.2	1.0	1.1	1.3	1.2	1.1	1.2	0.9	0.9
2007	0.9	0.8	1.0	1.1	1.2	1.2	1.3	1.2	1.1	1.2	1.0	0.9
2008	0.9	0.8	0.8	0.9	1.3	1.2	1.3	1.3	1.2	1.2	1.0	0.9
2009	0.9	0.8	0.8	1.0	1.3	1.2	1.3	1.2	1.2	0.8	1.1	0.9
2010	0.9	0.8	0.8	1.0	1.1	1.2	1.2	1.1	1.1	1.0	0.9	0.9
2011	0.9	0.8	0.8	1.0	1.2	1.1	1.3	1.3	1.3	1.2	0.9	0.8
2012	0.9	0.8	0.9	1.2	1.2	1.2	1.2	1.3	1.3	0.8	0.7	1.0
2013	0.9	0.8	0.8	0.7	1.3	1.1	1.2	1.3	1.2	1.2	0.9	0.9

2.3.2 MODIS recorded level 3 annual land-cover product (MCD12Q1). Its land-cover classification product following the University of Maryland (UMD) scheme has been used in this study. According to the UMD classification system, land-cover is classified into 14 different classes (Table 2.2). The land-cover dataset is available at a 500m spatial resolution for 2002-2012 at an annual time-step. Land-cover for the year 2013 is assumed to be the same as that of the year 2012 since annual land-cover data for this year was not available from the MODIS data repository. This is a reasonable

assumption since land-cover in the past has not changed drastically for this region at annual time-steps (see Figure 2.2 for instance).

2.3.3 The National Aeronautics and Space Administration (NASA) Shuttle Radar Topographic Mission (SRTM) elevation product. This data has a spatial resolution of 90 m.

2.3.4 The 3-hourly surface temperature estimates for the period 2013-2100 made by three GCMs from the Coupled Model Inter-comparison Project-Phase 5 (CMIP5) experiment has been used. A list of the selected GCMs is provided in Table 2.3. Climate model projections based on two Representative Concentration Scenarios (RCPs): RCP2.6 and RCP8.5 are used.

2.3.5 CMIP5 based daily near surface air temperature estimates for the period 2013-2100 are used in this study. Climate models and RCPs as specified before are considered.

2.3.6 The daily maximum air temperature (t_{max}) and minimum air temperature (t_{min}) ANUSPLIN data. These data are prepared by applying a thin plate smoothing spline surface fitting on the daily Environment Canada climate station observations (Hutchinson et al. 2009; Hopkinson et al. 2011). In this study, ANUSPLIN data for the period 2010-2013, encompassing the Canadian land-mass has been used.

2.3.7 North American Regional Reanalysis (NARR) 3-hourly surface temperature data for the period 2006-2013 is used. These data are produced by running a high resolution physical model (NCEP Eta model) together with the Regional Data Assimilation System (RDAS). Therefore although these data are model based, they are temporally and spatially synchronised with the observation records (Mesinger et al. 2006).

Table 2.2. Land-cover classes as identified by the UMD classification system. MODIS land-cover classification codes as well as abbreviations used for different land-cover classes in this study are also provided.

S.No	UMD classes	LHZ classes	UMD-LHZ classes
1	Water (W)	Water (W)	W
2	Evergreen Needleleaf Forest (ENF)	Pasture (G)	ENF-PR
3	Evergreen Broadleaf Forest (EBF)	Crop (C)	EBF-PR
4	Deciduous Needleleaf Forest (DNF)	Urban land (UB)	DNF-PR
5	Deciduous Broadleaf Forest (DBF)	Primary land (PR)	DBF-PR
6	Mixed Forest (MF)	Secondary land (SC)	MF-PR
7	Closed Shrublands (CS)		CS-PR
8	Open Shrublands (OS)		OS-PR
9	Woody Savannas (WS)		WS-PR
10	Savannas (S)		S-PR
11	Grasslands (G)		C
12	Croplands (C)		UB
13	Urban and Built-up (UB)		BSV-PR
14	Barren or Sparsely Vegetated (BSV)		ENF-SC
15			EBF-SC
16			DNF-SC
17			DBF-SC
18			MF-SC
19			CS-SC
20			OS-SC
21			WS-SC
22			S-SC
23			BSV-SC
24			G

Table 2.3. List of GCMs considered for analysis in this study.

GCM	Model	Resolution	Source
1	IAP-FGOALS	$1.66° \times 2.81°$	Institute of Atmospheric Physics, Chinese Academy of Sciences, China
2	MRI-CGCM3	$1.08° \times 2.16°$	Meteorological Research Institute, Japan
3	NorESM1-M	$2° \times 2°$	Norwegian Climate Centre, Norway

2.3.8 The land-use harmonization (LHZ) data (Hurtt et al. 2011) provide continuous land-use scenarios that smoothly connect past reconstructions of land-use based on HYDE data (Klein Goldewijk et al. 2001; Klein Goldewijk et al. 2010; Klein Goldewijk et al. 2011) with future projections based on the Integrated Assessment Model (IAM) implementations of different RCPs. In this study, version "LUHa_u2.v1" of the LHZ data product spanning the period: 2013 to 2100 has been used. In this version, six different land-use classes as provided in Table 2.2 have been included. For each LHZ grid, estimates of the total grid-area that translates from one LHZ class to another is provided at annual timesteps. The LHZ data is available at 0.5° x 0.5° resolution. The LHZ data also comes with a map (referred as "fnf" in this study) which can help identify whether or not a particular LHZ grid-cell is potentially forested or not. This map is based on the potential biomass density (pbd) outputs of the Miami model at each LHZ grid-cell. Both primary and secondary land in the LHZ data can be forested or non-forested. To help identify whether or not the secondary land within a LHZ grid can be called as forest, maps of secondary mean biomass density (sbd) and secondary mean age (sma) are also provided. As recommended in Hurtt et al. (2011), any vegetation with biomass density greater than $2kgC/m^2$, are considered as forest in this study.

2.4 Methods

In this section, methods and models used in this study have been described.

2.4.1 Bias Corrected Spatial Disaggregation (BCSD) downscaling approach: The BCSD downscaling method was proposed by Wood et al. (2004). The method consists of two major steps. In the first step, bias in GCM data is corrected using quantile mapping

approach. Both climate model and observed data are de-trended. The de-trended climate data are then used to construct a Cumulative Distribution Function (CDF) for the model $C_{raw,model}$ and an inverse cumulative function for the observed data C_{obs}^{-1}. The bias corrected model data $T_{bc,model}$ can then be found as:

$$T_{bc,model}(t) = C_{obs}^{-1}\left[C_{raw,model}\left\{ T_{raw,model}(t)\right\}\right] \tag{2.1}$$

Where, t represents the time step of analysis. Subscripts *raw* and *bc* are associated with the raw and bias-corrected version of climate model data respectively. In this study the bias correction is performed at monthly time-steps. Gridded observed climate data required for performing bias-correction using above approach is obtained by averaging the remotely sensed data at all pixels located within the target GCM grid. Secondly, bias corrected climate model data are spatially disaggregated by bilinearly interpolating, and then applying a fine-resolution spatial anomaly pattern from the observations. Traditionally this method has been applied on a monthly time scale. However more recently it has also been applied using daily time-steps as well (Abatzoglou et al. 2012; Thrasher et al. 2012). In this study the spatial anomaly pattern has been derived at daily scale. This spatial anomaly pattern is calibrated over the calibration period and is held fixed to downscale climate model data for the validation period.

2.4.2 Physical Scaling (SP) approach: Downscaling by SP approach is performed by establishing a multiple linear regression model. In this model, remotely sensed surface temperature data are considered as the response variable and bi-linearly interpolated

climate model data, elevation and land-cover data are considered as explanatory variables. The model can be mathematically expressed as:

$$ST_{rs,p} = \beta_0 + \beta_{mod} \times ST_{mod,p} + \beta_E \times E_p + \beta_{LC} \times LC_p + \varepsilon \tag{2.2}$$

Where, *ST* denotes surface temperature, *E* denotes elevation (masl), *LC* denotes the categorical land-cover variable which can take UMD land-cover class codes (provided in Table 2.2) as input values, ε denotes the error term associated with the regression model. Subscript *rs* and *mod* describe if the data is remotely sensed or model based, respectively. Further subscript *p* signifies that the regression is performed at a pixel level. Regression coefficients: β_{mod}, β_E and β_{LC} denote slopes associated with model based data, elevation and each land-cover class respectively. A separate model is developed for each of the 12 months. Further models are developed for day and night separately. Therefore in total, 24 different models are calibrated and used for prediction in this study.

2.4.3 Land-cover downscaling approach: Since LHZ data is available at 0.5° x 0.5° resolution, it needs to be downscaled to 500 m spatial resolution before it can be used to downscale future surface temperatures. Further land-use transition data are available in LHZ classes (Table 2.2). They need to be associated with appropriate UMD land-cover classes. Therefore we need to downscale as well as reconfigure land-use transition data into 500 m UMD land-cover classification before they can be used to model future land-cover projections. Following steps are performed for doing the same:

Setting up intermediate land-cover classes

First a set of intermediate land-cover classes are created. These are referred to as UMD-LHZ classes in this study and are summarized in Table 2.2. These classes act as a link between the UMD and LHZ classes and all subsequent land-cover change analysis are performed in this classification scheme. The classes that are common in both UMD and LHZ classification schemes such as: W, C, G and UB are included directly in the UMD-LHZ classification scheme. Other UMD-LHZ classes are created by merging the UMD and LHZ classes where first part of the UMD-LHZ class name comes from the associated UMD class and the second part comes from the associated LHZ class.

Preparation of UMD-LHZ data for the base year: 2012

Land-cover downscaling process starts with the preparation of land-cover data in UMD-LHZ classification scheme for the baseline year: 2012. For doing so, MODIS land-cover data for the year 2012 (in UMD land-cover classification) is analyzed. Pixels associated with UMD classes: W, C, G and UB are directly translated to respective UMD-LHZ classes since these classes are common in both classification schemes. Pixels belonging to UMD classes: ENF, EBF, DNF, DBF, MF, CS, OS, WS, S and BSV are further distributed into primary and secondary sub-classes to obtain pixels belonging to UMD-LHZ classes 2-10 and 13-23 respectively (Table 2.2). Hurtt et al. (2011) defines primary land as areas that have not been impacted by human activities in the past whereas secondary land as areas that have been impacted by human activities in the past and are recovering. We use the same definition to segregate pixels belonging to a particular UMD class into primary and secondary categories. For

doing so, land-cover data for above mentioned classes are analyzed over the period 2002-2012. Pixels found to have transitioned to these classes within this period are taken as secondary pixels while the rest are considered as primary pixels. By the end of this step, land-cover data for the year 2012 is obtained in UMD-LHZ classification scheme.

Locating transitioning pixels

Annual land-use transition data (in LHZ classification) for the year 2012 is extracted at all LHZ grids located within the study region. This includes the number of pixels that will transition from one land-use class to other. Within each HRZ grid cell, pixels belonging to a particular LHZ class (which can encompass one or more UMD-LHZ classes) that are most likely to transit from one class to the other are located using a distance based rule. Pixels with a particular land-use class that are closely grouped together are considered to be more resistant to change than the ones that are isolated from each other. Such neighborhood based rules have been incorporated in previous land-cover change studies (West et al. 2014; Verburg et al. 2004a) to account for spatial auto-correlation in a) environmental features that govern landscape development and b) land-use expansion being most dominant around a similar land-use area (Verburg et al. 2004b) and have been adopted in this study as well.

Finding destination UMD-LHZ class of transitioning pixels

Destination LHZ class of a transitioned pixel is extracted from the LHZ data. It is associated with a UMD-LHZ class using the following rules:

- If the destination LHZ class is W/C/G/UB, then UMD-LHZ class considered is W/C/G/UB respectively since these classes are common between LHZ and UMD-LHZ classes.

- If the destination LHZ class is PR and fnf (described in section 2.3.8) = 1, then the output land-cover is of primary-forest type. The UMD based forest cover class of the transitioned pixels is obtained using the model defined in 2.4.4. The obtained forest cover class is used to obtain corresponding UMD-LHZ class associated with the transitioning pixel.

- If the destination LHZ class is PR and fnf = 0, then the output land-cover class is BSV-PR.

- If the destination LHZ class is SC, fnf = 1 and sbd >= 2, then the output land-cover class is of secondary-forest type. The UMD based forest cover class of the transitioned pixels will be obtained using the model defined in 2.4.4. The obtained forest cover class is used to obtain corresponding UMD-LHZ class associated with the transitioning pixel.

- If the destination LHZ class is SC, fnf = 1 and sbd < 2, then the output land-cover class is of secondary-nonforest type. The output land-cover class is BSV-SC.

- If the destination LHZ class is SC, fnf = 0 and sbd >= 2, then the output land-cover class is of secondary-nonforest type. The output land-cover class is BSV-SC.

- If the destination LHZ class is SC, fnf = 0 and sbd < 2, then the output land-cover class is of secondary-nonforest type. The output land-cover class is BSV-SC.

Generation of land-cover data for subsequent years

Land-use transition data for subsequent years: 2013-2100 are extracted and steps 2.4.3 are repeated at annual time-steps to obtain future annual land-cover in UMD-LHZ classification.

Regrouping land-cover classes from UMD-LHZ to UMD classification scheme

Future land-cover obtained in UMD-LHZ classification are regrouped to get land-cover in UMD classification schemes by merging classes that were segregated before. Pixels belonging to classes: W, G, C, UB are transitioned directly from UMD-LHZ to UMD class.

2.4.4 Forest-cover type model: The UMD forest cover type associated with a LHZ forest land-use class is obtained by using a multinomial logit model calibrated on the climate type associated with different forest cover types across Canada. The regression model is calibrated using MODIS land-cover data for the year 2012 and yearly averaged ANUSPLIN precipitation and temperature data for the period 2010-2013. The reason behind considering more than one year of climate data is to ensure that yearly fluctuations in climate are ignored during the model development process. The multinomial logit model describing forest–cover type can be mathematically expressed as:

$$\log(\frac{p_{FT,i}}{p_{FT,ref}}) = \alpha_0 + \alpha_1 \times T_p + \alpha_2 \times P_p + \varepsilon \tag{2.3}$$

Where, $p_{FT,i}$ denote the probability of membership of the pixel in the ith forest cover class, $p_{FT,ref}$ denote the probability of membership in reference forest cover class. The forest-cover classes considered to fit the model are: ENF, EBF, DNF, DBF, MF, CS, OS, WS and S. T denotes the ANUSPLIN temperature data associated with the MODIS pixel, P denotes the ANUSPLIN precipitation data associated with the MODIS pixel and ε represents the error term associated with the model. Further subscript p denotes that the data are a pixel level data.

Above model is used to obtain probabilities associated with different UMD forest cover types given a set of precipitation and temperature conditions associated with the pixel of interest. In this analysis future precipitation and air temperature projections made by the three GCMs are used to obtain future land-cover specific probabilities. Thereafter future forest-cover type (in UMD classification) is obtained by selecting the most probable forest-cover class among all forest cover classes present within a LHZ grid-cell.

2.5 Results and discussion

2.5.1 Evaluation and comparison of downscaling approaches

NARR based surface temperature (ST) data are downscaled using five different approaches: 1) BCSD model, 2) Bias-correction and Bilinear Interpolation (BCBI) method, 3) Bilinear Interpolation (BI) method, 4) SP method, and 5) SP method neglecting land-cover as a predictor (SP-LC). Approaches 1, 2 and 4 have been explained in section 2.4. The BCBI method involves bias-correcting reanalysis ST data using quantile mapping approach (explained before) and then performing bilinear

34

interpolation to get local ST data. In the BI method local ST values are obtained by bilinear interpolation of raw reanalysis data. In SP-LC method, land-cover is ignored as a predictor of ST in the SP model equation (equation 2.2). In other words, ST is modelled considering reanalysis based ST and elevation as predictors.

The calibration of models is performed over the period: 2006-2010 while the validation is performed on the year 2013. The reason behind selecting these calibration and validation periods is to test the performance of these methodologies in a climate that is distinctly different from the calibration period. By an analysis of the historical remotely sensed surface temperature data, the period 2006-2010 was found to be significantly different from the period 2010-2013. The total numbers of data pixels available for the validation year (2013) are found to be 337120. Before using any remotely sensed data product, their quality assessment files are referred, and only pixels with reliable data are selected for analysis. In the case of surface temperature, pixels which are associated with <1 K of error are deemed as reliable while land-cover pixels which are deemed as of "good quality" in the remotely sensed datasets are considered reliable. Since remotely sensed surface temperature, land-cover and elevation data are available at different spatial resolutions, all datasets were resampled to a common resolution level for analysis, one that is associated with the land-cover data (i.e. 500 m). Further, model based datasets are temporally interpolated using the nearest available hourly values to obtain hourly data at an instant at which no data is available. For instance, in order to obtain day-time surface temperature datasets for the Terra satellite (which crosses the study region close to 10:30 AM), data recorded at 9 AM and 12 PM are averaged to estimate the associated grid-value at that time.

The performance of models towards downscaling NARR ST outputs is evaluated by comparing the downscaled outputs with the remotely sensed ST data over the validation period. Two metrics are chosen to evaluate the models: 1) Root Mean Squared Error (RMSE) in surface temperatures ($RMSE_{st}$) and 2) RMSE in mean land-cover specific surface temperatures ($RMSE_{st-lc}$). In the tables 2.4 and 2.5, the $RMSE_{st}$ and $RMSE_{st-lc}$ values for above mentioned downscaling approaches are presented. The performances of methodologies are evaluated for snow-free days (sf-day), snow-free nights (sf-night), snow-covered days (sc-day) and snow-covered nights (sc-night) separately. The downscaling method performing best at each timeline is highlighted in orange. Overall SP method is found to be the best performing method ($RMSE_{st} = 5$ K and $RMSE_{st-lc} = 2$ K) followed by SP-LC method ($RMSE_{st} = 5$ K and $RMSE_{st-lc} = 2$ K), followed by BCBI method ($RMSE_{st} = 7$ K and $RMSE_{st-lc} = 3$ K), followed by BI method ($RMSE_{st} = 11$ K and $RMSE_{st-lc} = 10$ K), followed by BCSD method ($RMSE_{st} = 13$ K and $RMSE_{st-lc} = 11$ K).

Table 2.4. Root Mean Squared Error (RMSE) in surface temperatures as predicted by the BCSD, SP, SP-LC, BI and BCBI models. RMSE values are shown for different scenarios considered for analysis, in day and night, and for snow-free (sf) and snow-covered (sc) conditions. Best performing model has been highlighted in orange for each timeline and scenario analyzed.

Scaling	Timeline			
	sf-day	*sf-night*	*sc-day*	*sc-night*
BCSD	14.8	12.5	10.1	15.0
SP	5.0	4.4	5.0	5.4
SP-LC	5.6	4.5	5.1	5.5
BI	11.9	13.9	7.2	12.2
BCBI	8.7	6.7	5.0	6.2

Table 2.5. Root Mean Squared Error (RMSE) in surface temperatures as predicted by the BCSD, SP, SP-LC, BI and BCBI models. RMSE values are shown for different scenarios considered for analysis, in day and night, and for snow-free (sf) and snow-covered (sc) conditions. Best performing model has been highlighted in orange for each timeline and scenario analyzed.

Scaling	Timeline			
	sf-day	*sf-night*	*sc-day*	*sc-night*
BCSD	13.1	11.2	7.6	12.7
SP	1.6	1.3	2.2	2.3
SP-LC	2.8	1.5	2.3	2.2
BI	10.9	13.2	4.8	10.8
BCBI	4.1	3.4	2.6	3.7

2.5.2 SP model performance assessment

Over the validation period, SP model is found to perform slightly better in the nighttime ($RMSE_{st}$ = 4.9 K and $RMSE_{st-lc}$ = 1.8 K) than in daytime ($RMSE_{st}$ = 5.0 K and $RMSE_{st-lc}$ = 1.9 K) and in snow-free months ($RMSE_{st}$ = 4.7 K and $RMSE_{st-lc}$ = 1.4 K) than in snow-covered months ($RMSE_{st}$ = 5.2 K and $RMSE_{st-lc}$ = 2.2 K). A reason for a lower performance in the daytime and snow-covered days can be attributed to a lower density of remotely sensed data in the daytime than in the nighttime, and during snow-covered months than the snow-free months as discussed before in section 2.2. Lower model performance in the snow-covered months can also occur because snow cover and snow depth have not been considered as predictor variables in this study. Snow cover is a very important factor which influences winter time climate dynamics in Saskatchewan and hence ignoring it is expected to have detrimental effects on the model performance.

The effectiveness of proposed downscaling methodology towards capturing variability in mean land-cover temperatures is evaluated. The Pearson correlation coefficient (r) between the modelled and observed mean land-cover surface temperature curves for different timelines are shown in Table 2.6. It can be noted that r value is greater than 0.8 for all timelines considered for analysis. Further Figure 2.3 presents the mean land-cover surface temperature curves for timelines showing highest (Snow-free, Aqua, Night) and lowest (Snow-covered, Aqua, Night) correlations. It can be seen that the model is able to simulate mean land-cover temperature variability very well across all land-cover classes present within the study region. It is also found that model performs better in low elevation regions than high elevation regions. This can be seen from Figure B1 where elevation and RMSE distribution across the study region is presented. It is clear from the figure that model performance varies systematically with elevation distribution.

Table 2.6. Pearson correlation coefficient (r) between modelled and observed mean land-cover surface temperature curves for all timelines and snow-cover states considered for analysis.

Satellite	Time	Snow	Correlation (r)
Aqua	Day	sf	0.90
		sc	0.90
	Night	sf	0.98
		sc	0.81
Terra	Day	sf	0.89
		sc	0.97
	Night	sf	0.93
		sc	0.95
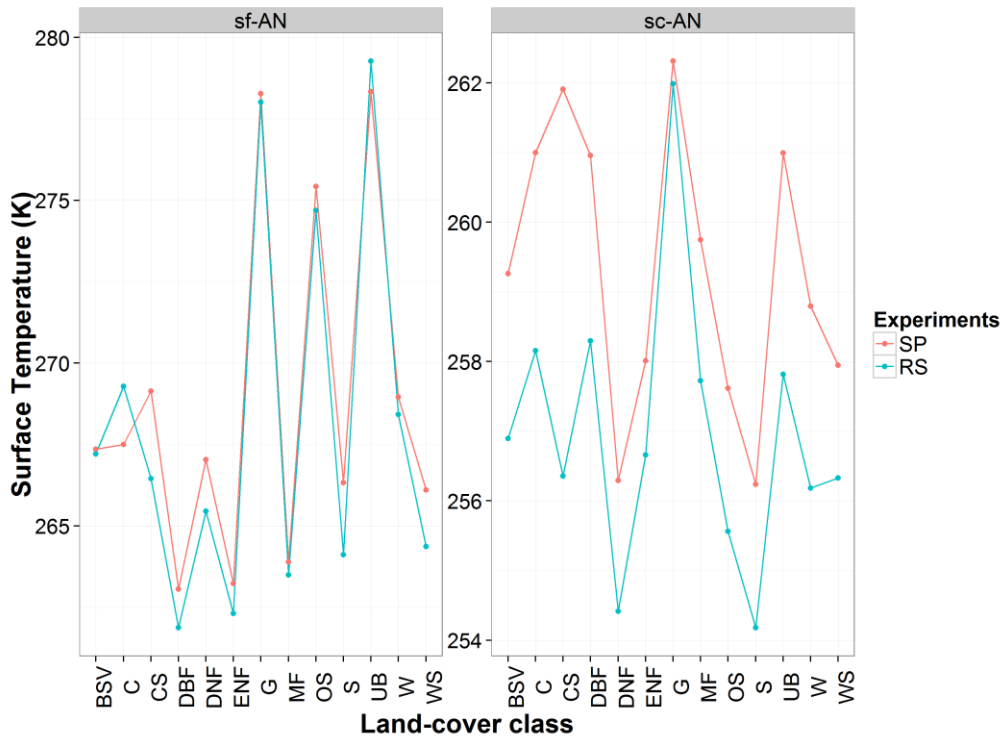
Figure 2.3. Modelled and remotely sensed mean land-cover temperatures corresponding to scenario: mri-cgcm3-rcp26 for timelines: a) snow-covered (night) and b) snow-free (night). The red and blue lines represent the observed and simulated values respectively.

2.5.3 Future land-cover projection

Land-cover for the period 2013-2100 is generated using the methodology explained before. Expected land-cover trajectories across the study region within this period for both RCPs considered for analysis are shown in Figure 2.4. Under the RCP 2.6 scenario, a reconfiguration in the tree cover types is projected. An analysis of the area fraction occupied by different tree cover types suggests a gradual transition from forest classes: ENF, EBF, DNF, DBF, MF and CS to OS and S classes. The total area encompassed by the OS land-cover class is projected to increase to almost 50 times the current area while the S land-cover class is projected to encompass 17 times more area.

Minor increases in G, WS and C land-cover classes as well as slight decreases in the BSV land-cover class area are also observed. Under the RCP 8.5 scenario, most striking development over the 21st century is found to be the transition from all land-cover classes to BSV land-cover class. The total BSV land-cover area almost doubles at the expense of other land-cover classes. The area occupied by every other land-cover class either decreases or remains constant over the 21st century. Most significant decreases are observed for the MF land-cover class however other tree types: ENF, DNF, DBF, MF, CS, WS and S also decrease. A slight decrease in G land-cover area and increase in C land-cover area is also projected.

Furthermore it is found that the differences in temperature as projected by different climate models do not impact future tree-cover distribution significantly. This is found by comparing land-cover projections corresponding to different climate models under a single emission scenario. It is found that the relative ranks of different forest-cover classes (in terms of probability of occurrence) did not vary significantly for different climate models under a fixed emission scenario.
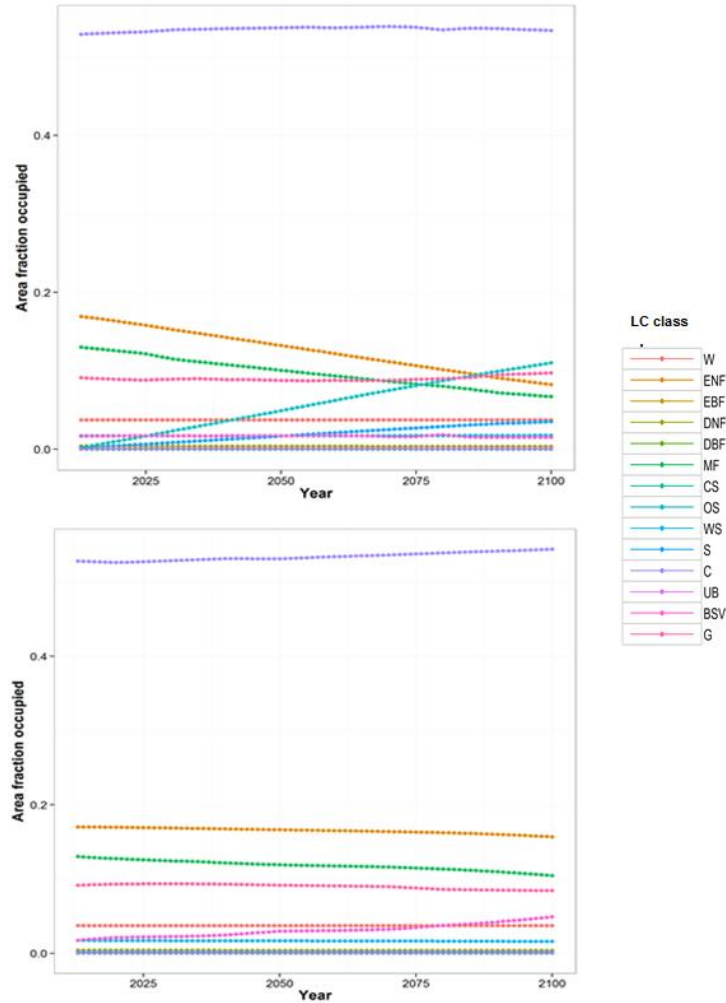
Figure 2.4. Annual land-cover projections across the study region over the period: 2013-2100. Data corresponding to scenario mri-cgcm3-rcp2.6 (top) and mri-cgcm3-rcp8.5 (bottom) have been presented.

2.5.4 Future surface temperature projections

The SP model is recalibrated over the baseline period 2006-2013 and is used to downscale surface temperatures projections from three CMIP5 models for the future period 2014-2100. Future land-cover, elevation and bi-linearly interpolated GCM data at each pixel as projected under different emission scenarios are used to make

downscaled future surface temperature projections. Model parameters are kept constant between the baseline and future timelines.

Downscaled surface temperature shows an overall increasing trend across the 21st century as illustrated in Figure 2.5. In this figure, yearly mean surface temperatures averaged across all climate models are shown. The rate of increase in surface temperatures is higher in the case of scenario: RCP8.5 (0.04 K/year) as compared to the scenario: RCP2.6 (0.003 K/year). Under the RCP2.6 scenario, surface temperatures are projected to increase to reach a maximum value by the mid of 21st century and then decrease thereafter. On the other hand, under the RCP8.5 scenario, continuous increase in surface temperatures is obtained over the 21st century.
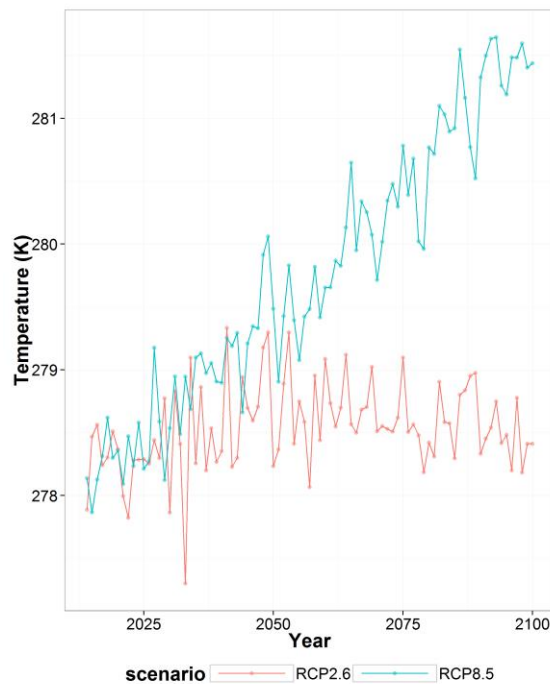
Figure 2.5. Future surface temperature trends associated with the two emission scenarios. Climate projections have been averaged across all GCMs to obtain the trends.

The sensitivity of the projected changes is explored with the help of change factors (CF) which are defined as the difference in mean surface temperature between baseline and the year: 2100. Figure 2.6 presents CFs for different land-covers for daytime and nighttime. It can be seen that the CFs differ for different land-cover classes. In the daytime most significant increase in temperature are recorded for grasslands (2 K), followed by croplands (1.4 K), followed by BSV (1.3 K), followed by UB (1 K), and followed by forests (0.9 K). On the other hand, in the nighttime most significant change in temperature are obtained for forest classes (1.7 K), followed by C (1.6 K), followed by BSV (0.9 K), followed by UB (0.8 K) and followed by G (-0.1 K). Overall largest changes are observed for C, followed by forest-cover classes, followed by BSV, followed by UB, and followed by G. Significant variability in the projected changes are noticed among forested land-cover classes. Overall EBF is found to be show smallest changes (0.9 K) while CS are found to be associated with the largest changes (1.7 K).
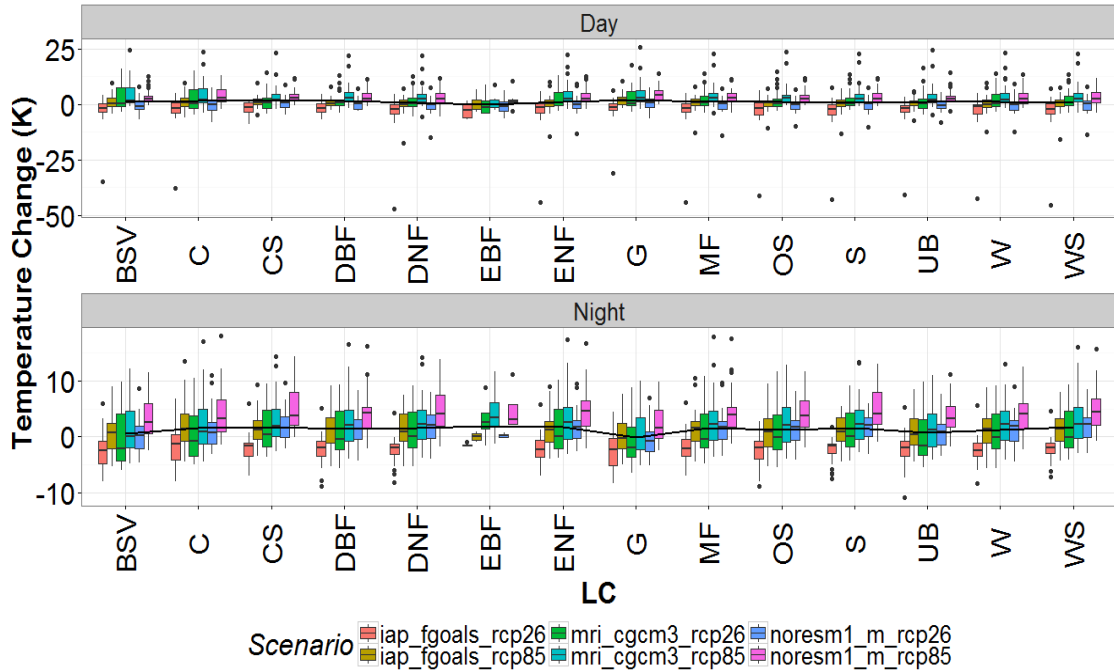
Figure 2.6. Temperature change as projected for different land-covers considered in this study in day and night for all scenarios considered in this study. The black trend-line denotes the average variations from all scenarios.

The variation of CF values with elevation is also explored. Table 2.7 presents the rates of change in CF with elevation for all models, scenarios, snow-states and time of the day. Mean land-cover CFs are subtracted from the raw CFs to obtain elevation specific CFs. A linear regression is thereafter performed with CF anomaly as the predictant variable and elevation as predictor variable and the rate of change in CF with elevation is estimated. It is found that the CFs increase with elevation in the snow-free months (rate = 7e-04 K/m) whereas they decrease with elevation in the snow-covered months (rate = -3e-03 K/m). This suggests that higher altitude regions may experience larger increases in temperatures than the low-lying regions during the snow-free months, whereas low lying regions may experience larger increase in temperature than high

altitude regions during the snow-covered regions. This can also be noted from the CF vs. elevation plots presented in Figure 2.7. Here the results have been presented for a representative model (noresm1-m), scenario (RCP2.6) and time of the day (TN).

Among different climate models considered, largest changes (2.4 K) are projected by the climate model: mri-cgcm3, followed by noresm1-m (2.2 K) and followed by iap-fgoals (-0.9 K). Further larger changes are projected for RCP8.5 scenario (2.7 K) than RCP2.6 scenario (-0.22 K). Further difference in changes are obtained for snow-covered months (2.1 K) than the snow-free months (0.1 K), and in the nighttime (1.3 K) than in the daytime (1.1 K). Lastly a comparison of different sources of uncertainty is performed by analyzing the magnitude of changes projected by the three GCMs, two emission scenarios, two snow-cover states, four time of the day and 14 land-cover classes considered in this study. The results are presented in Figure 2.8 where the uncertainty magnitude for above mentioned sources is shown. As it can be observed from Figure 2.8, GCM is found to be the most important source of uncertainty (uncertainty range = 3.2 K), followed by the choice of RCP (uncertainty range = 2.9 K), followed by the snow-cover state of the area (uncertainty range = 2.0 K), followed by the land-cover class (uncertainty range = 0.9 K), and followed by the time of the day (uncertainty range = 0.2 K).
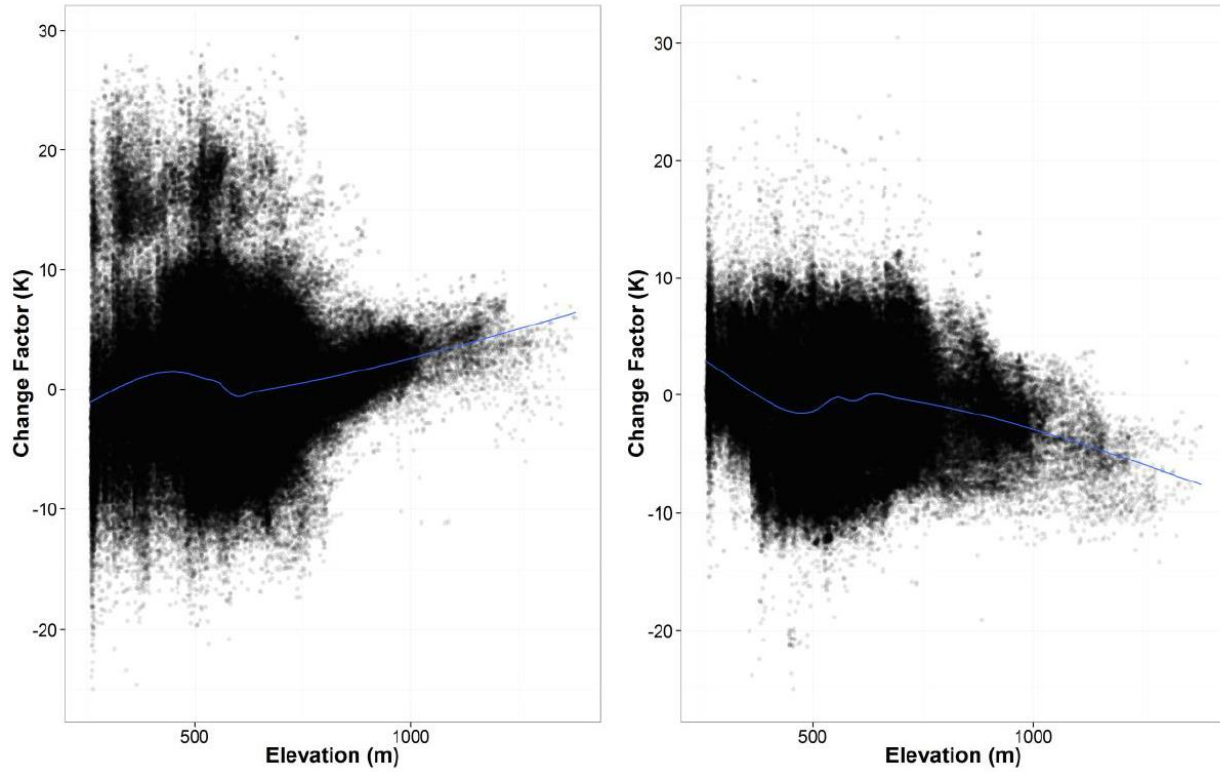
Figure 2.7. Variation of change factor (CF) with elevation for snow-free (left) and snow-covered (right) months. Blue line shows the smoothed fitted line by obtaining a Generalized Additive Model (GAM) fit between CF and elevation. This representative result is shown for model: noresm1-m, scenario: RCP2.6, time: AN.

Table 2.7. Variation of Change Factors (CF) with elevation. Results are presented for all models, scenario, snow-cover state, time of the day considered for analysis.

Model	Scenario	Snow	Time	Change in ST with elevation (K/m)
mri-cgcm3	RCP2.6	sf	day	2e-03
	RCP8.5			-7e-05
	RCP2.6	sc		-6e-03
	RCP8.5			-8e-03
	RCP2.6	sf	night	1e-03
	RCP8.5			-1e-03
	RCP2.6	sc		1e-04

	RCP8.5			-2e-03
iap-fgoals	RCP2.6	sf	day	2e-03
	RCP8.5			-7e-05
	RCP2.6	sc		-6e-03
	RCP8.5			-8e-03
	RCP2.6	sf	night	1e-03
	RCP8.5			-1e-03
	RCP2.6	sc		1e-04
	RCP8.5			-2e-03
noresm1-m	RCP2.6	sf	day	2e-03
	RCP8.5			-7e-05
	RCP2.6	sc		-6e-03
	RCP8.5			-8e-03
	RCP2.6	sf	night	1e-03
	RCP8.5			-1e-03
	RCP2.6	sc		1e-04
	RCP8.5			-2e-03
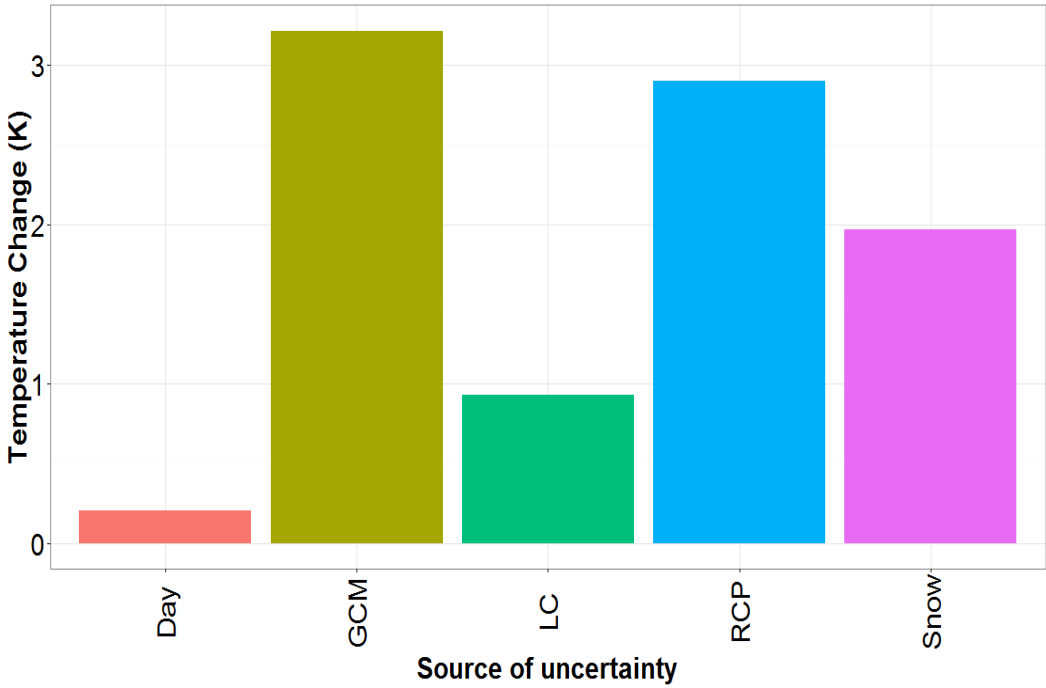
Figure 2.8. A comparison of the magnitude of uncertainty associated with all five sources of uncertainty considered in this study.

2.6 Conclusions

A physically scaling (SP) model has been introduced in this study which can be used to downscale climate model based surface temperature datasets under non-stationary future conditions. The model is based on a hypothesis that local scale climate can be modelled using large scale climate and land-cover, elevation characteristics of the location of interest. From the results presented in this study, it can concluded that above mentioned hypothesis is true and hence the proposed model can be used to downscale future temperatures under changing climatic and land-cover conditions as simulated by the GCMs (Stocker et al. 2013) and Integrated Assessment Models (Hurtt et al. 2011).

This study also presents a case study of the proposed method on the southern Saskatchewan region of Canada. From the analysis of future temperature projections many interesting results are obtained. For instance, it is found that land-cover and snow-cover properties of a particular location play an important role in shaping its response to changes in large scale climate. Our analysis shows that their influence at local and regional scales is comparable and even bigger than that contributed by differences in climate models and emission scenarios. Further considerable differences in the projected changes are obtained at a diurnal scale. Lastly it is also shown that the projected temperature changes vary systematically with elevation. It is found that during the snow-free months, the lower elevation or relatively flat regions are more resistant to temperature increases than the higher elevation regions. On the other hand, during the snow-covered months higher elevation regions are more resistant to temperature increases than the low lying regions. These results highlight the importance of considering the geophysical properties of the location of interest and temporal scale of analysis while making future climate projections at local and regional scales.

The current work can be extended in many different directions. In the proposed SP model, pixels belonging to a particular land-cover class are considered to vary only with changes in elevation. However land-cover and elevation properties of the surrounding pixels also play an important role in deciding the climatology of a location. The proposed model can therefore be improved in future by incorporating these neighborhood characteristics into SP model formulation. Further, it is important to ascertain an appropriate spatial scale at which SP model calibration should be performed to get the most accurate yet stable regression parameters. Finally the model can be used to downscale temperature in other regions of the world which have more complicated physiography than the region that has been selected for this study.

References

Abatzoglou JT, Brown TJ (2012) A comparison of statistical downscaling methods suited for wildfire applications. International Journal of Climatology 32: 772-780.

Adachi SA, Kimura F, Kusaka H, Inoue T, Ueda H (2012) Comparison of the impact of global climate changes and urbanization on summertime future climate in the Tokyo Metropolitan Area. Journal of Applied Meteorology and Climatology 51:1441–1454.

Argueso D, Evans JP, Fita L, Bormann, KJ (2014) Temperature response to future urbanization and climate change. Climate Dynamics 42(7-8):2183-2199.

Das S, Simonovic SP (2012) Assessment of Uncertainty in Flood Flows under Climate Change Impacts in the Upper Thames River Basin, Canada. British Journal of Environment & Climate Change 2(4): 318–338.

Encyclopedia of Saskatchewan. University of Regina. http://esask.uregina.ca/entry/climate.html. Accessed May 7 2015.

Ezber Y, Sen OL, Kindap T, Karaca M (2007) Climatic effects of urbanization in Istanbul: a statistical and modeling analysis. International Journal of Climatology 27: 667-679.

Fall S, Niyogi D, Gluhovsky A, et al. (2010) Impacts of land use land cover on temperature trends over the continental United States: assessment using the North American Regional Reanalysis. Int. J. Climatol. 30:1980–1993.

Fowler HJ, Blenkinsop S, Tebaldi C (2007) Review: Linking climate change modeling to impacts studies: Recent advances in downscaling techniques for hydrological modeling. Int. J. Climatol. 27:147–178.

Gaur A, Simonovic SP (2015) Towards reducing climate change impact assessment process uncertainty. Environmental Processes 10.1007/s40710-015-0070-x.

Georgescu M, et al. (2013) Summer-time climate impacts of projected megapolitan expansion in Arizona. Nature Climate Change 3:37-41.

Hale RC, Gallo KP, Owen TW, Loveland TR (2006) Land use/land cover change effects on temperature trends at U.S. Climate Normals stations. Geophysical Research Letters 33: L11703.

Hale RC, Gallo KP, Loveland TR (2008) Influences of specific land use/land cover conversions on climatological normals of near-surface temperature. Journal of Geophysical Research 113: D14113.

Hamdi R, de Vyver HV, de Troch R, Termonia P (2014) Assessment of three dynamical urban climate downscaling methods: Brussels's future urban heat island under an A1B emission scenario. Int. J. Climatol. 34: 978–999.

Hopkinson RF, et al. (2011) Impact of aligning climatological day on gridding daily maximum-minimum temperature and precipitation over Canada. Journal of Applied Meteorology and Climatology 50: 1654:1665.

Hu Y, Jia G (2010) Influence of land use change on urban heat island derived from multi-sensor data. International Journal of Climatology 30: 1382-1395.

Hurtt GC, et al. (2011) Harmonization of land-use scenarios for the period 1500–2100: 600 years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands. Clim. Change 109(1-2): 117–161.

Hutchinson MF, et al. (2009) Development and testing of Canada-wide interpolated spatial models of daily minimum-maximum temperature and precipitation for 1961-2003. Journal of Applied Meteorology and Climatology 48: 725:741.

IPCC. 2013: Summary for Policymakers. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker,T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Ke X, Ma E, Yuan Y (2014) Scenario Simulation of the Influence of Land Use Change on the Regional Temperature in a Rapidly Urbanizing Region: A Case Study in Southern-Jiangsu, China. Advances in Meterology 159724:1-12.

King LM, McLeod AI, Simonovic SP (2014) Simulation of historical temperatures using a multi-site, multivariate block resampling algorithm with perturbation. Hydrol. Process. 28: 905–912.

Klein Goldewijk K (2001) Estimating global land use change over the past 300 years: The HYDE database. Global Biogeochemical Cycles 15(2): 417–433.

Klein Goldewijk K, Beusen A, Janssen P (2010) Long term dynamic modeling of global population and built-up area in a spatially explicit way: HYDE 3.1. The Holocene 20: 565–573.

Klein Goldewijk K, Beusen A, van Drecht G, de Vos M (2011) The HYDE 3.1 spatially explicit database of human induced land use change over the past 12,000 years. Global Ecol Biogeogr 20: 73–86.

Kusaka H, Hara M, Takane Y (2012) Urban Climate Projection by the WRF Model at 3-km Horizontal Grid Increment: Dynamical Downscaling and Predicting Heat Stress in the 2070's August for Tokyo, Osaka, and Nagoya Metropolises. Journal of the Meteorological Society of Japan 90B: 47-63.

Limonsu A, Viguie V, Daniel M, Masson V (2015) Vulnerability to heat waves: Impact of urban expansion scenarios on urban heat island and heat stress in Paris (France). Urban Climate 14: 586-605.

Lowry WP (1977) Empirical estimation of the urban effects on climate: A problem analysis. J. Appl. Meteor. 16:129–135.

Mesinger F, DiMego G, Kalnay E, Mitchell K, and Coauthors (2006) North American Regional Reanalysis. Bulletin of the American Meteorological Society 87: 343–360.

Oke TR (1982) The energetic basis of the urban heat island. Quart. J. Roy. Meteor. Soc. 108:1–24.

Oke TR (1987) Boundary Layer Climates. 2nd ed. Routledge, 435 pp.

Pielke RA, Pitman A, Niyogi D, Mahmood R, McAlpine C, Hossain F, Goldewijk KK, Nair U, Betts R, Fall S, Reichstein M, Kabat P, Noblet N (2011) Land use/land cover changes and climate: modeling analysis and observational evidence. WIREs Clim Change 2: 828-850.

Roth M, Chow WTL (2012) A historical review and assessment of urban heat island research in Singapore. Singapore Journal of Tropical Geography 33: 381-397.

Qiao Z, Guangjin T, Lixiao Z, and Xu X (2014) Influences of Urban Expansion on Urban Heat Island in Beijing during 1989–2010. Advances in Meteorology 187169. doi:10.1155/2014/187169.

Salathé EP (2003) Comparison of various precipitation downscaling methods for the simulation of streamflow in a rainshadow river basin. International Journal of Climatology 23: 887–901.

Salathé EP (2005) Downscaling simulations of future global climate with application to hydrologic modeling. International Journal of Climatology 25: 419–436.

Schoof JT (2013) Statistical downscaling in climatology, Geogr. Compass 7:249–265.

Srivastav R, Simonovic SP (2014) Multi-site, multivariate weather generator using maximum entropy bootstrap. Climate Dynamics doi: 10.1007/s00382-014-2157-x.

Stewart ID (2000) Influence of meteorological conditions on the intensity and form of the urban heat island effect in Regina. The Canadian Geographer / Le Géographe canadien 44: 271–285.

Teutschbein C, Seibert J (2012) Bias correction of regional climate model simulations for hydrological climate-change impact studies: review and evaluation of different methods. J. Hydrol. 456:12–29.

Thrasher B, Maurer EP, McKellar C, Duffy PB (2012) Technical Note: Bias correcting climate model simulated daily temperature extremes with quantile mapping. Hydrol. Earth Sys. Sci. 16: 3309-3314.

Verburg PH, de Nijs TCM, van Eck JR, Visser H, de Jong K (2004a) A method to analyse neighbourhood characteristics of land use patterns. Computers, Environment and Urban Systems 28 (6): 667-690.

Verburg PH, Schot PP, Dijst MJ, Veldkamp A (2004b) Land use change modelling: current practice and research priorities. Geojournal 61: 309-324.

Wang T, Hamann A, Spittlehouse DL, Murdock TQ (2011) ClimateWNA – high-resolution spatial climate data for western North America. Journal of Applied Meteorology and Climatology 51:16–29.

West TO, Page YL, huang M, Wolf J, Thomson AM (2014) Downscaling global land cover projections from an integrated assessment model for use in regional analyses: results and evaluation for the US from 2005 to 2095.

Widmann M, Bretherton CS, Salathé EP (2003) Statistical precipitation downscaling over the Northwestern United States using numerically simulated precipitation as a predictor. Journal of Climate 16: 799–816.

Wilks DS (2012) Stochastic weather generators for climate-change downscaling, part II: multivariable and spatially coherent multisite downscaling. WIREs Clim Change 3: 267–278.

Wood AW, Leung LR, Sridhar V, Lettenmaier DP (2004) Hydrologic implications of dynamical

and statistical approaches to downscaling climate model outputs. Climatic Change 62: 189–

216.

CHAPTER 3: Extension of SP method and its application towards downscaling climate model based near surface air temperature

3.1 Introduction

Climate models are mathematical representation of the globe and can simulate complex physical processes occurring within the earth's climate system. They are therefore perfectly placed to simulate large scale climatic response to increasing greenhouse gases in the earth system. For local or regional scale climate change impact assessment studies these large scale climatic changes need to be transferred to an appropriate local or regional scale. This process of extraction of local or regional scale information from large scale climatic projections is referred to as downscaling (Stocker et al. 2013).

Two broad streams of downscaling methodologies exist in the climate change literature: 1) Dynamic downscaling and 2) Statistical downscaling. Dynamic downscaling involves using Regional Climate Models (RCMs), which are essentially high resolution mesoscale models that can model climatic processes operating at spatial scales much smaller than those resolved by the GCMs (Xue et al. 2014). RCMs use boundary conditions provided by the GCMs and distribute it across the study region in a physically based way. Statistical downscaling methods aim to perform a similar task and employ statistical methods for doing so. Here large scale climate model data is linked with observed point location data using statistical methods (Fowler et al. 2007; Schoof et al. 2013). Both dynamic and statistical methods of downscaling have advantages and disadvantages. While dynamic approaches are physically based, they are computationally expensive (Xue et al. 2014). On the other hand, statistical approaches are computationally inexpensive but are not physically based.

With an aim at providing physically representative downscaled products for climate models, Gaur and Simonovic (2016) proposed a Physical Scaling (SP) based statistical downscaling approach. Since the proposed approach is statistical in nature, it is computationally inexpensive and the downscaled outputs are physically representative, if not physically based. Scaling approaches have been used to downscale climate model data in the past. These approaches model local scale climate based on large scale value of the same climate variable (Schoof et al. 2013). For instance Wang et al. (2011) performed bi-linear interpolation with lapse rate adjustments to downscale air temperature data across western North America. Salathe (2003) used scaling based approaches to downscale precipitation in the Yakima River basin (USA) and found them to be effective in capturing precipitation dynamics across the catchment. Wood et al. (2004) used three different scaling methods to downscale climate model and RCM generated gridded precipitation and temperature data. They found that scaling based downscaling methods are able to capture the observed hydrometeorologic variability in their outputs.

In Gaur and Simonovic (2016) SP model is calibrated by formulating a linear regression model with bilinearly interpolated NARR surface temperature data, Moderate-resolution imaging spectroradiometer (MODIS) based land-cover, and elevation as predictor variables and MODIS based surface temperature data as predictant variable. The calibrated model is thereafter used to downscale future projected GCM surface temperature data across the southern Saskatchewan region in Canada. The three predictors used in SP method are selected based on the recommendations made by Lowry (1977) and others (Oke 1982; Fall et al. 2010; Hale et al. 2006; 2008; Argueso et al. 2014; Kusaka et al. 2012; 2014; Kishtawal et al. 2010; Rao et al. 2004; Li et al. 2011; Shepherd 2005; Efe 2014; Lin et al. 2013) who

found that elevation, land-cover and large scale climate shape locally observed climate. The approach is found to be able to simulate surface temperatures and mean land-cover specific surface temperatures across the study region significantly better than a state-of-the-art statistical downscaling methodology: Bias Correction Statistical Downscaling (BCSD) method (Wood et al. 2004).

In this study the validity of SP method is tested on another important climate variable: near surface air temperature. Two approaches towards downscaling air temperature have been proposed: direct and indirect. In the direct approach air temperature data is downscaled directly using SP method. On the other hand in the indirect approach, SP model is first used to downscale climate model based surface temperature data. The downscaled surface temperature data is thereafter used to estimate air temperature using another statistical model that links surface temperature to air temperature (referred as ST → AT model hereafter). Estimation of air temperature from surface temperatures has been performed in many studies in the past. Many statistical functions like linear regression (Stathopoulou et al. 2006), random forests (Xu et al. 2014), optimization techniques (Benali et al. 2012), M5 method (Emamifar et al. 2013), kriging method (Anderson 2002) have been used to establish relationship between surface and air temperatures. Further a range of predictors in addition to surface temperatures have been used for instance land-cover (White-Newsome et al. 2013; Xu et al. 2012), julian day and day length (Benali et al. 2012), solar radiation (Emamifar et al. 2013), Normalised Difference Vegetation Index (Goetz et al. 2000 and Stisen et al. 2007), solar zenith angle (Vogt et al. 1997).

Several models following the direct and indirect approaches are evaluated in this study. Model ensemble is prepared by considering different functional forms, neighborhood scales,

and predictor variables. Further a sensitivity analysis is performed to quantify and compare the impact of the choice of functional form, neighborhood scale and downscaling approach on the predicted future air temperature. Rest of the paper is organized as follows. The study region and datasets used are provided in sections 3.2 and 3.3 respectively. This is followed by a description of the models and methods used in this study in section 3.4. A discussion on results obtained is provided in section 3.5 followed by conclusions in section 3.6.

3.2 Study region

The region selected for analysis is the southern Saskatchewan region in Canada (see Figure 3.1). The area is land-locked and is in abundance of small lakes and rivers. The region is characterized by large topographic variability and by the presence of different land-cover classes. The elevation across the study region varies from 240 masl to 1389 masl while all land-cover classes identified in the University of Maryland (UMD) classification scheme (see Table 3.1) are present. The region has forested land-cover in the north and cropland, grassland areas in the south. Overall, cropland and forests are two major land-cover classes occupying the study region accounting for close to 80% of the total area. Two major urban centers: Saskatoon and Regina are present within the study region.

The climate of Saskatchewan is continental and is characterized by its extremes. Large fluctuations in temperature (up to 65ºC) can be observed within a year owing to its land-locked position in the North American land-mass. Due to this the region heats up as well as cools down quickly. An important climatic feature of the region is frequent clear skies and sunny conditions. Majority of precipitation that Saskatchewan receives occurs during the summers due to the passing of mid-latitude cyclones over the region. Wintertime

precipitation occurs as snow and due to sustained below zero temperatures accumulated snow-pack has a major influence on the climatology of the region (Encyclopedia of Saskatchewan 2015).
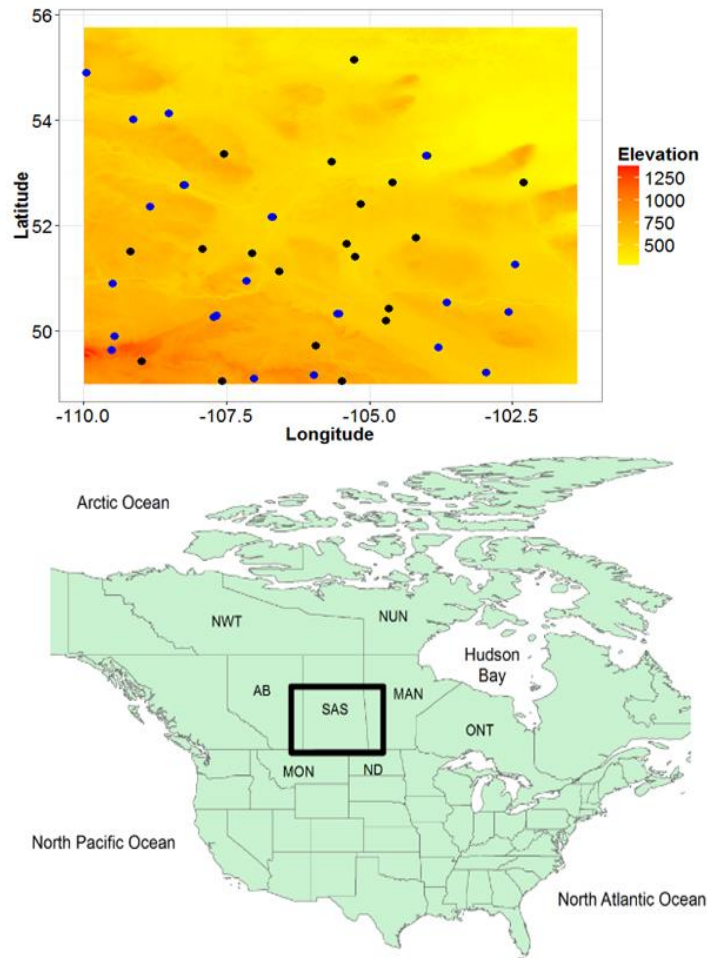
Figure 3.1. Location and physiography of the study region considered for analysis. The black and blue dots in the top figure show the location of calibration and validation air temperature recording stations respectively.

3.3 Data used

SRTM elevation data: The National Aeronautics and Space Administration (NASA) Shuttle Radar Topographic Mission (SRTM) elevation product is used in this study. This data has a spatial resolution of 90 m.

MODIS land-cover and surface temperature data: MODIS recorded level 3 annual land-cover product (MCD12Q1) in the UMD classification scheme has been used in this study. A list of land-cover classes identified in the UMD classification scheme is provided in Table 3.1. The annual land-cover data, available at a 500 m spatial resolution is obtained for the period 2006-2012. Land-cover data for the year 2013 is not available from MODIS data repository therefore it is assumed to be the same as that of the year 2012. This is a reasonable assumption since land-cover in the study region has not changed drastically at annual time-steps in the past (Gaur and Simonovic 2016). MODIS recorded surface temperature level 3 products from Aqua (MYD11A1) and Terra (MOD11A1) are also used in this study. Terra passes equator at around 10:30 AM/PM while Aqua passes at around 1:30 PM/AM. Day-time as well as night-time surface temperature are collected for the period 2006-2013 from both the satellites. The percentage distribution of the remotely sensed surface temperature data over the study period is provided in Table 3.2. Before using any remotely sensed product their quality assessment files are referred and only pixels with reliable data are selected for analysis. In the case of surface temperature, the pixels which are associated with <1 K of error are deemed as reliable while land-cover pixels which are deemed as of "good quality" in the remotely sensed datasets are considered reliable.

Table 3.1. Land-cover classes as identified in the UMD classification system. Abbreviation used

for each land-cover class is also provided within the brackets.

S.No	UMD classes
1	Water (W)
2	Evergreen Needleleaf Forest (ENF)
3	Evergreen Broadleaf Forest (EBF)
4	Deciduous Needleleaf Forest (DNF)
5	Deciduous Broadleaf Forest (DBF)
6	Mixed Forest (MF)
7	Closed Shrublands (CS)
8	Open Shrublands (OS)
9	Woody Savannas (WS)
10	Savannas (S)
11	Grasslands (G)
12	Croplands (C)
13	Urban and Built-up (UB)
14	Barren or Sparsely Vegetated (BSV)

Table 3.2. Distribution (%) of the remotely sensed surface temperature data over the

period 2006-2013.

Year	Month											
	J	*F*	*M*	*A*	*M*	*J*	*J*	*A*	*S*	*O*	*N*	*D*
2006	0.8	0.7	0.8	1.2	1.0	1.1	1.3	1.2	1.1	1.2	0.9	0.9
2007	0.9	0.8	1.0	1.1	1.2	1.2	1.3	1.2	1.1	1.2	1.0	0.9
2008	0.9	0.8	0.8	0.9	1.3	1.2	1.3	1.3	1.2	1.2	1.0	0.9
2009	0.9	0.8	0.8	1.0	1.3	1.2	1.3	1.2	1.2	0.8	1.1	0.9
2010	0.9	0.8	0.8	1.0	1.1	1.2	1.2	1.1	1.1	1.0	0.9	0.9
2011	0.9	0.8	0.8	1.0	1.2	1.1	1.3	1.3	1.3	1.2	0.9	0.8
2012	0.9	0.8	0.9	1.2	1.2	1.2	1.2	1.3	1.3	0.8	0.7	1.0
2013	0.9	0.8	0.8	0.7	1.3	1.1	1.2	1.3	1.2	1.2	0.9	0.9

Recorded hourly air temperature: In total 71 hourly air temperature recording stations are found to be located within the study region. Historically observed hourly air temperature data for the period 2006 to 2013 are collected from the Environment Canada database at these stations. Out of them 52 stations are found to have satisfactory data length. A list of these stations is provided in Table 3.3. All recording stations are found to be associated with cropland or grassland land-cover class based on MODIS land-cover data.

Table 3.3. List of calibration and validation stations selected for analysis.

S.No	Station name	Latitude	Longitude	Elevation
	Calibration			
1	Rosetown East	51.57	-107.92	586.00
2	Last Mountain Cs	51.42	-105.25	497.00
3	Bratt's Lake Climate	50.20	-104.71	580.00
4	Wynyard (Aut)	51.77	-104.20	560.10
5	Nipawin	53.33	-104.00	371.90
6	Assiniboia Airport	49.73	-105.95	725.50
7	Hudson Bay(Aut)	52.82	-102.32	358.10
8	Pilger	52.42	-105.15	552.00
9	Prince Albert A	53.22	-105.67	428.20
10	Outlook Pfra	51.48	-107.05	541.00
11	North Battleford	52.77	-108.25	548.00
12	Coronach Spc	49.05	-105.48	756.00
13	Watrous East	51.67	-105.40	525.60
14	Melfort	52.82	-104.60	490.00
15	Elbow Cs	51.13	-106.58	595.00
16	Kindersley A	51.52	-109.18	693.70
17	Meadow Lake A	54.13	-108.52	480.70
18	North Battleford Rcs	52.77	-108.26	548.00
19	Yorkton	51.26	-102.46	498.40
20	Eastend Cypress (Aut)	49.44	-108.99	1059.00
21	Spiritwood West	53.37	-107.55	584.30
22	Yorkton	51.26	-102.46	498.30
23	La Ronge A	55.15	-105.27	379.20
24	Regina Int'l A	50.43	-104.67	577.60
25	Regina Rcs	50.43	-104.67	577.30
26	Saskatoon Intl A	52.17	-106.70	504.10

27	Val Marie Southeast	49.06	-107.59	796.00
Validation				
28	Broadview	50.37	-102.57	599.80
29	Estevan	49.22	-102.97	580.60
30	Estevan A	49.22	-102.97	580.30
31	Indian Head Cda	50.55	-103.65	579.10
32	Loon Lake Rcs	54.02	-109.14	545.60
33	Lucky Lake	50.95	-107.15	664.70
34	Meadow Lake	54.13	-108.52	481.00
35	Moose Jaw A	50.33	-105.57	576.70
36	Moose Jaw Cs	50.33	-105.54	577.00
37	Nipawin	53.33	-104.02	371.90
38	Nipawin	53.33	-104.01	371.90
39	North Battleford	52.77	-108.24	548.30
40	Rockglen (Aut)	49.17	-105.98	917.00
41	Saskatoon Rcs	52.17	-106.72	504.10
42	Scott Cda	52.36	-108.83	659.60
43	Swift Current	50.29	-107.69	816.90
44	Swift Current A	50.30	-107.68	816.90
45	Swift Current Cda	50.27	-107.73	825.00
46	Weyburn	49.70	-103.80	588.60
47	Yorkton	51.26	-102.46	498.30
48	Cypress Hills Park	49.64	-109.51	1259.00
49	Jimmy Lake Awos	54.91	-109.96	637.10
50	Leader Airport	50.91	-109.50	675.50
51	Mankota	49.10	-107.02	830.00
52	Maple Creek	49.90	-109.47	766.70

NARR data: NARR 3-hourly data for air temperature, surface temperature, and atmospheric variables: wind-speed at 10m, upward longwave radiation flux, upward shortwave radiation flux, low cloud area fraction, medium cloud area fraction, high cloud area fraction and specific humidity is obtained for the period 2006-2013. The NARR data has an approximate spatial resolution of 32 Km (Mesinger et al. 2006).

Future land-cover projections: Future land-cover projections in UMD classification scheme for the period: 2081-2100 is obtained from Gaur and Simonovic (2016). In Gaur and

Simonovic (2016) yearly future land-cover data for the period 2014-2100 is obtained by associating future land-use projections provided in Hurtt et al. (2011) with MODIS land-cover classes. The obtained future land-cover data has a spatial resolution of 500 m.

 GCM air temperature data: Future 3-hourly air temperature, surface temperature, total cloud fraction, eastward wind, northward wind, specific humidity, surface downwelling longwave radiation, surface downwelling shortwave radiation, surface upwelling longwave radiation, surface upwelling shortwave radiation data from a General Circulation Model (GCM): FGOALS-s2 (Qing et al. 2012) is collected for an emission scenario: Representative Concentration Pathway (RCP) 8.5 for the period 2081-2100.

3.4 Approaches, methods and models used

In this section models, methods and approaches considered in this study are described. A list of the models considered is provided in Table 3.4.

3.4.1 Downscaling approaches

Two different approaches are adopted for downscaling model based air temperature data: 1) direct and 2) indirect. The direct approach involves a one-step procedure of the application of SP method (and its extensions) using recorded and climate model based air temperature data. Since recorded data is used as a predictant variable in this approach, model calibration can only be performed over land-cover pixels associated with the recording stations. Model predictions therefore can only be made at land-cover pixels associated with these land-cover classes. The indirect approach involves two steps. In the first step SP method is used to downscale model based surface temperature data. Remotely sensed surface temperature data are used as predictant variable while

model based surface temperature data (along with other predictor variables) is used as predictor variable. Since remotely sensed data is used as a predictant variable, model calibration (and prediction) can be performed at all land-cover class pixels that are associated with the remotely sensed data. In the second step ST → AT is used to estimate air temperature from the downscaled surface temperature data. Since the density of remotely sensed data is much higher than the density of recording stations, the indirect approach is expected to yield a higher density of downscaled air temperature data than the direct method.

Table 3.4. Models evaluated in this study.

S.No	Approach	Model	Method	Functional form	Predictors (ST →AT model)
1		SP_lm	SP	LR	-
2		SP_qr	SP	QR	-
3		SP_gam	SP	GAM	-
4	Direct	SPS3x3_lm	SPS	LR	-
5		SPS5x5_lm	SPS	LR	-
6		SPS7x7_lm	SPS	LR	-
7		SPS9x9_lm	SPS	LR	-
8		SP_lm_ST	SP	LR	ST
9		SP_qr_ST	SP	QR	ST
10		SP_gam_ST	SP	GAM	ST
11		SP_lm_ST.LC	SP	LR	ST, LC
12		SP_qr_ST.LC	SP	QR	ST, LC
13		SP_gam_ST.LC	SP	GAM	ST, LC
14	Indirect	SP_lm_ST.LC.AVs	SP	LR	ST, LC, AVs
15		SP_qr_ST.LC.AVs	SP	QR	ST, LC, AVs
16		SP_gam_ST.LC.AVs	SP	GAM	ST, LC, AVs
17		SPS3x3_lm_ST.LC.AVs	SPS	LR	ST, LC, AVs
18		SPS5x5_lm_ST.LC.AVs	SPS	LR	ST, LC, AVs
19		SPS7x7_lm_ST.LC.AVs	SPS	LR	ST, LC, AVs
20		SPS9x9_lm_ST.LC.AVs	SPS	LR	ST, LC, AVs

3.4.2 SP method and its extensions

SP method

Downscaling by SP method is performed by forming a multiple linear regression model with observed climate data as predictant variable and bilinearly interpolated climate model data, elevation and land-cover as predictor variables. The SP method formulation for the downscaling of climate model based air temperatures can be mathematically expressed as:

$$CV_{obs} = \beta_0 + \beta_1 \times CV_{mod} + \beta_2 \times E_p + \beta_3 \times LC_p + \varepsilon \qquad (3.1)$$

Where, CV denotes the climate variable of interest, E denotes the elevation (masl), LC denotes the categorical land-cover variable, β denote the regression parameters and ε denotes the error term associated with the regression model. Subscript *obs* and *mod* describe if the climatic data is observed or model based, respectively. Subscript p indicates that the data used is a pixel scale data. In Table 3.4, models with SP method are denoted with a prefix "SP".

SP method with Surrounding pixel information (SPS method)

The SP method is modified to incorporate land-cover and elevation configuration surrounding the pixel of interest. The mathematical formulation of the SPS method can be expressed as:

$$CV_{obs} = \beta_0 + \beta_1 \times CV_{mod} + \beta_2 \times E_p + \beta_3 \times LC_p + Fr_{W,s} + ... + Fr_{BSV,s} + R_{E,s} + \varepsilon \qquad (3.2)$$

In the SPS method, additional neighborhood land-cover pixel information of a reference pixel is incorporated by adding predictors that convey the fraction of surrounding area that is occupied by each UMD class. For instance in equation 3.2 additional predictors $Fr_{W,s}$, ...$Fr_{BSV,s}$ represent the fraction of the total area surrounding the reference pixel by *Water,….Barren and Sparsely Vegetated* land-cover classes. The value of each of these predictors is between 0 and 1 and they add up across all land-cover classes to give a value of 1. Neighborhood elevation information is incorporated by including a predictor $R_{E,s}$ which represents the ratio between reference pixel elevation and mean elevation of pixels surrounding the reference pixel. In all additional predictors, the subscript *s* denotes that the predictors are calculated at a certain neighborhood scale. In this study the analysis is performed at four neighborhood scales: 3x3, 5x5, 7x7 and 9x9 as adopted in some studies in the past (White and Engelen 2000; Verberg et al. 2004). Configuration of neighborhood scales considered in this analysis is shown in Figure 3.2. The reference pixel is shown in red. Neighborhood pixels encompassed in 3x3, 5x5, 7x7 and 9x9 neighborhood scale are shown in light red, light green, light blue and grey respectively. Areas encompassed in higher neighborhood scales are inclusive of smaller neighborhood scales. This means that neighborhood area of 5x5 scale will encompass the area associated with 3x3 neighborhood scale plus the light green area. In Table 3.4, models with SP method are denoted with a prefix "SPS". Models calibrated at 3x3, 5x5, 7x7 and 9x9 neighborhood scales are referred as SPS3x3, SPS5x5, SPS7x7 and SPS9x9 respectively.
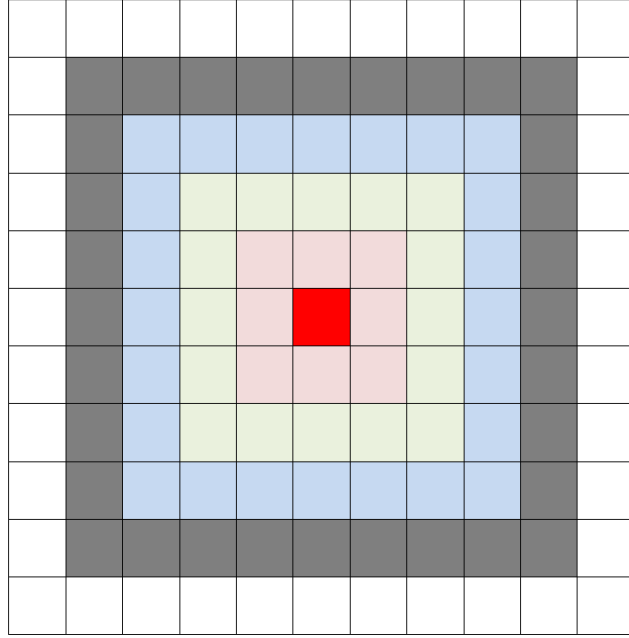
Figure 3.2. Neighborhood configurations considered in this study. In this figure reference pixel is shown in red and pixels encompassed in 3x3, 5x5, 7x7 and 9x9 neighborhood scale are shown in light red, light green, light blue and grey respectively. Areas encompassed in higher neighborhood scales are inclusive of the smaller neighborhood scales.

SP method with other regression functions

Apart from the linear regression model (LR) described above, two other regression functions are used in SP model formulation: 1) quantile regression (QR) and 2) generalized additive models (GAM). In Table 4, models using functional form: LR, QR and GAM are denoted with suffix: "lm", "qr" and "gam" respectively. The mathematical formulations of these models are provided in equations 3.3 and 3.4 respectively.

$$CV_{obs}(q) = \beta_0^{\,q} + \beta_1^{\,q} \times CV_{mod} + \beta_2^{\,q} \times E_p + \beta_3^{\,q} \times LC_p + \varepsilon^q \tag{3.3}$$

$$g(CV_{obs}) = \beta_0{}^q + f_1(CV_{\text{mod}}) + f_2(E_p) + f_3(LC_p) + \varepsilon \tag{3.4}$$

In equation 3.3, $CV_{obs}(q)$ indicates CV_{obs} values below quantile q. Further $\beta_0{}^q, \beta_1{}^q, \beta_2{}^q$

and $\beta_3{}^q$ are quantile specific parameters and ε^q is the quantile specific error

component of the model. In this study only the parameters associated with 0.5^{th} quantile

(median) is considered while making prediction using quantile regression. In equation

3.4, g is the link function and f_1, f_2, f_3 represent the non-parametric smoothed function

that is associated with model based data, elevation and land-cover respectively. In this

study, the smoothed function is fit using penalized likelihood maximization algorithm.

The penalized likelihood maximization algorithm is a variant of maximum likelihood

estimation algorithm and applies a tradeoff between model fit wiggliness and goodness

of fit by incorporating a penalty function (Wood 2000).

3.4.3 ST → AT models

Linear regression with surface temperature as predictor

The most basic model linking air temperature with surface temperature formulates a

linear regression relationship using surface temperature as predictor. The model

formulation can be mathematically expressed as:

$$AT = \beta_0 + \beta_1 \times ST + \varepsilon \tag{3.5}$$

Where, AT and ST denote air temperature and surface temperature respectively. In

Table 3.4, model using ST as predictor is denoted with a suffix: "ST".

Using other regression functions

In addition to LR model, QR and GAM based models are used to model air temperature from surface temperature. The models can be mathematically expressed as:

$$CV_{obs}(q) = \beta_0{}^q + \beta_1{}^q \times CV_{\text{mod}} + \varepsilon^q \tag{3.6}$$

$$g(CV_{obs}) = \beta_0 + f_1(CV_{\text{mod}}) + \varepsilon \tag{3.7}$$

The variables used in above equations are similar to those used in equations 3.3 and 3.4.

Using additional predictors

The basic models described in equations 3.5, 3.6 and 3.7 are extended by incorporating additional predictor variables such as land-cover as well as atmospheric variables (AVs): cloud-cover, specific humidity, upward longwave radiation flux, upward shortwave radiation flux and wind speed. In Table 3.4, models using land-cover and AVs as predictors are denoted with suffix: "LC" and "AVs" respectively.

3.5 Results and discussion

Each model listed in Table 3.4 is formulated separately for snow-covered (chosen as October to March) and snow-free months (chosen as April to September). The models are evaluated using two metrics: Root Mean Squared Error in predicted air temperature (RMSE-AT) and mean land-cover specific air temperature (RMSE-LC-AT). Further the spatial and temporal robustness of these models is tested by performing two sets of experiments:

E1 Test for temporal robustness: In this experiment, models are calibrated over the period 2006-2010 and validated over the period 2011-2013. Data from 52 stations located within the study region are used in this experiment.

E2 Test for spatial robustness: In this experiment, models are calibrated across 27 evenly distributed gauging stations located across the study region and validation across the rest 25 stations. Data for the entire period of study 2006-2013 is considered for analysis in this experiment.

The models considered in both direct and indirect approaches are found to performed better in the temporal robustness test than in the spatial robustness test. The RMSE-AT and RMSE-AT-LC found values associated with the direct approach for experiments: E1 and E2 are presented in Figures 3.3 and 3.4 respectively. The RMSE-AT and RMSE-AT-LC values from experiment E1 are found to be 0.06 K and 0.19 K respectively while from experiment E2 are found to be 0.13 K and 0.72 K respectively. The results from indirect downscaling approach for experiments E1 and E2 are presented in Figures 3.5 and 3.6 respectively. The RMSE-AT and RMSE-AT-LC values from experiment E1 are found to be 0.87 K and 0.93 K respectively while from experiment E2 are found to be 1.17 K and 2.16 K respectively.
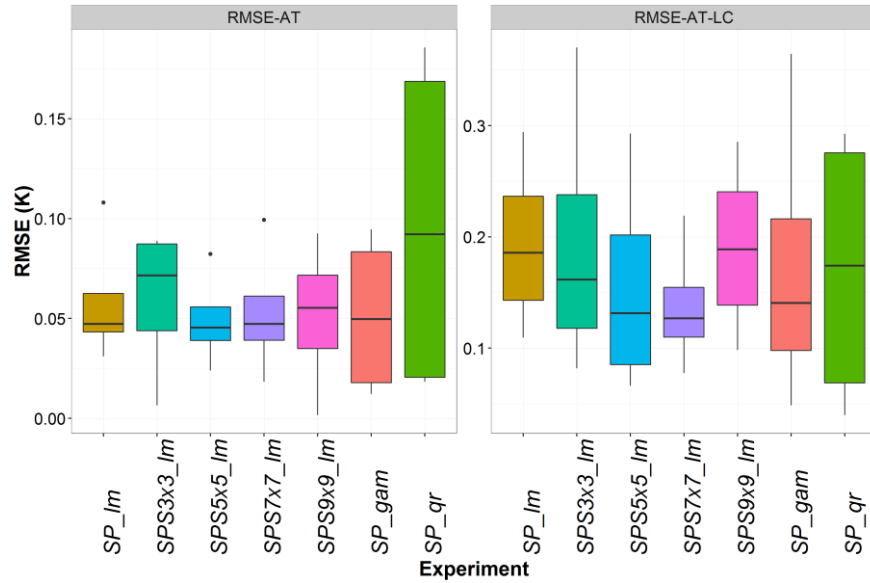
Figure 3.3. The RMSE associated the downscaled air temperature data (RMSE-AT) and mean land-cover air temperature (RMSE-AT-LC) using direct SP approach from the temporal robustness (E1) test.
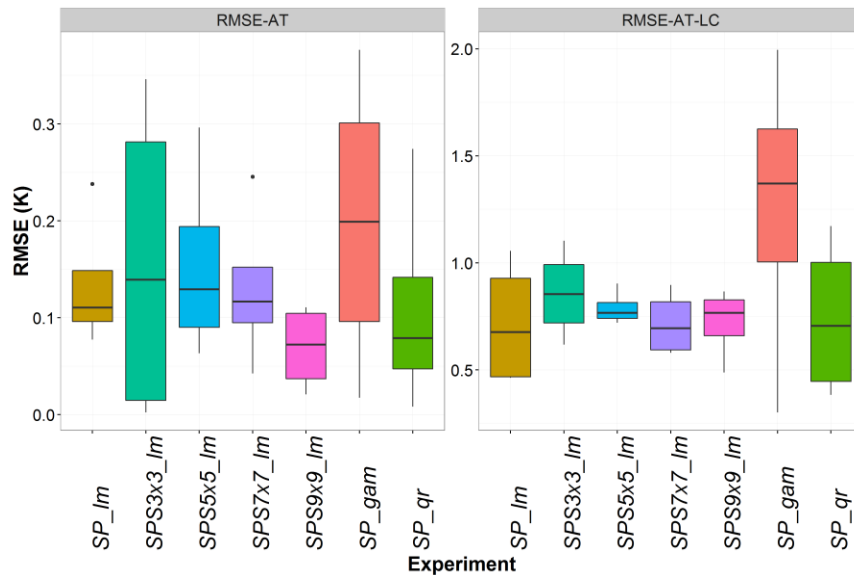
Figure 3.4. The RMSE associated the downscaled air temperature data (RMSE-AT) and mean land-cover air temperature (RMSE-AT-LC) using direct SP approach using direct SP approach from the spatial robustness (E2) test.
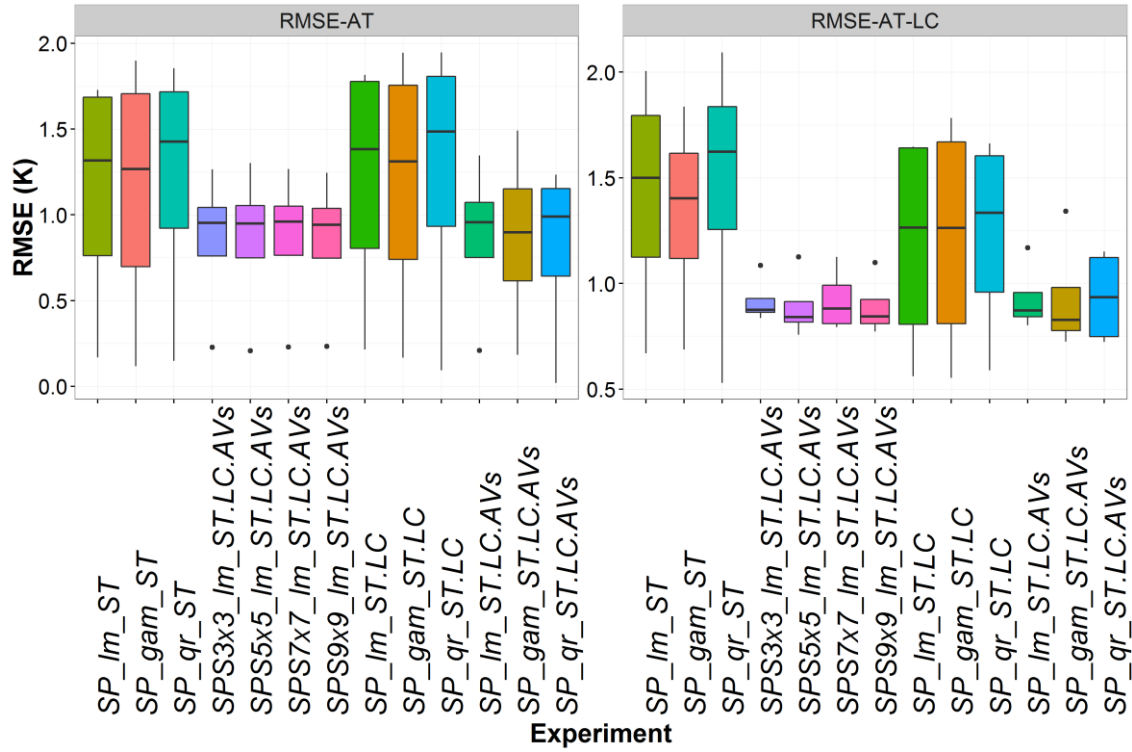
Figure 3.5. The RMSE associated the downscaled air temperature data using indirect SP

approach from the temporal robustness (E1) test.
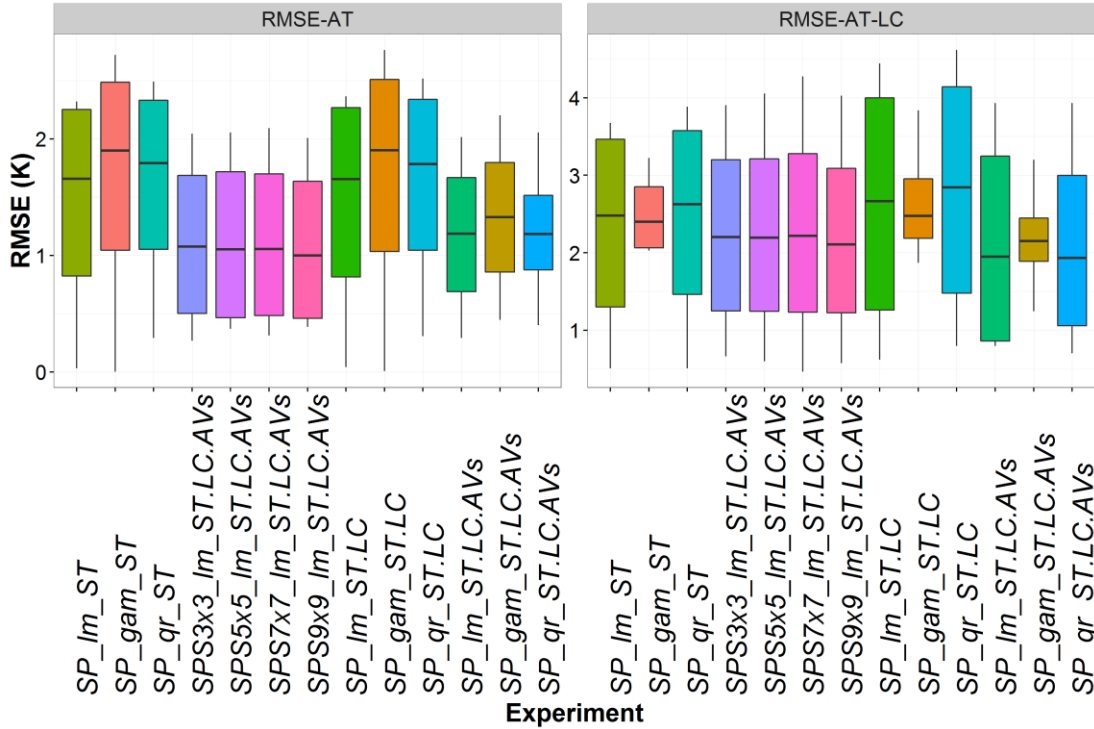
Figure 3.6. The RMSE associated the downscaled air temperature data using indirect SP approach from the spatial robustness (E2) test.

Among the two downscaling approaches, model performances are found to be better in the case of direct approach (RMSE-AT = 0.10 K; RMSE-AT-LC = 0.50 K) than the indirect approach (RMSE-AT = 1.17 K; RMSE-AT-LC = 1.73 K). Further superior model performance is obtained in the nighttime (RMSE-AT = 0.43 K; RMSE-AT-LC = 0.87 K) than in the daytime (RMSE-AT = 1.17 K; RMSE-AT-LC = 1.73 K).

Among the models considered under the direct approach (models 1-3 in Table 3.4), the *SP_gam* model is found to perform best (RMSE-AT = 0.05 K; RMSE-AT-LC = 0.17 K), followed by *SP_lm* (RMSE-AT = 0.06 K; RMSE-AT-LC = 0.19 K), followed by *SP_qr* (RMSE-AT = 0.10 K; RMSE-AT-LC = 0.17 K). Further by comparing model 1 with models 4-7 (in Table 3.4), it is found that the addition of neighborhood information at 3x3, 5x5, 7x7

and 9x9 neighborhood scale improves the performance of *SP_lm* model by 2%, 15%, 9% and 12% in terms of RMSE-AT and ~0%, 20%, 30% and 2% in terms of RMSE-AT-LC respectively.

Among the models considered in the indirect approach, models with functional forms: LR (models 8, 11, 14), GAM (models 10, 13, 16) and QR (models 9, 12, 15), models using GAM are found to perform best (RMSE-AT = 1.06 K; RMSE-AT-LC = 1.16 K), followed by LR (RMSE-AT = 1.07 K; RMSE-AT-LC = 1.18 K) and QR (RMSE-AT = 1.09 K; RMSE-AT-LC = 1.21 K). A comparison of models 17-20 with 14 suggests that the addition of neighborhood information increases the model performance. In terms of RMSE-AT an increase of 2%, 2%, 1% and 3% in model performance is observed for neighborhood scales: 3x3, 5x5, 7x7, and 9x9 respectively while in terms of RMSE-AT-LC an increase of 1%, 3%, 1% and 4% is observed. By comparing models with different predictors in the ST $\rightarrow$ AT model it is found that the addition of land-cover as an additional predictor to surface temperature leads to a decrease in the prediction accuracy by 4% in terms of RMSE-AT and an increase in prediction accuracy by 14% in terms of RMSE-AT-LC. The addition of AVs as predictors results in a significant improvement in the efficiency of models considered in the indirect approach. An increase of 30% and 23% in prediction accuracy is found in terms of RMSE-AT and RMSE-AT-LC respectively.

The sensitivity of future projections made by models listed in Table 3.4 is explored with reference to the usage of different functional forms, neighborhood scales and downscaling approaches considered in this study. Models provided in Table 3.4 are calibrated over the period 2006-2013 and used to make future air temperature projections for the period 2081-2100. Future air temperature projections made by models 1-7 during the snow-free months is

presented in figure 3.7 for timelines: Aqua-Day (AD), Terra-Day (TD), Aqua-Night (AN) and Terra-Night (TN). It can be seen from the figure that future projections are more sensitive to selected neighborhood scale than the functional form considered for analysis. Overall, it is found that over the period 2081-2100 mean air temperature varies by 0.1 K for the three functional forms however it varies by ~5 K between the four neighborhood scales considered in this study. This significant variation in the projections with different neighborhood scales is found to occur because of a variable response of neighborhood pixels on the reference pixel at different neighborhood scales. This can be observed from figure 8 where the rate of change in temperature with reference to increase in neighborhood land-cover fraction is provided for all land-cover classes for all neighborhood scale (NS) considered for analysis. Negligible and insignificant (at p=0.05) rates are shown in white whereas positive and negative rates are shown in red and green respectively. It can be seen the rates associated with each land-cover class vary significantly for different neighborhood scales highlighting the role that surrounding pixels play towards shaping significantly different temperature response at different neighborhood scales.
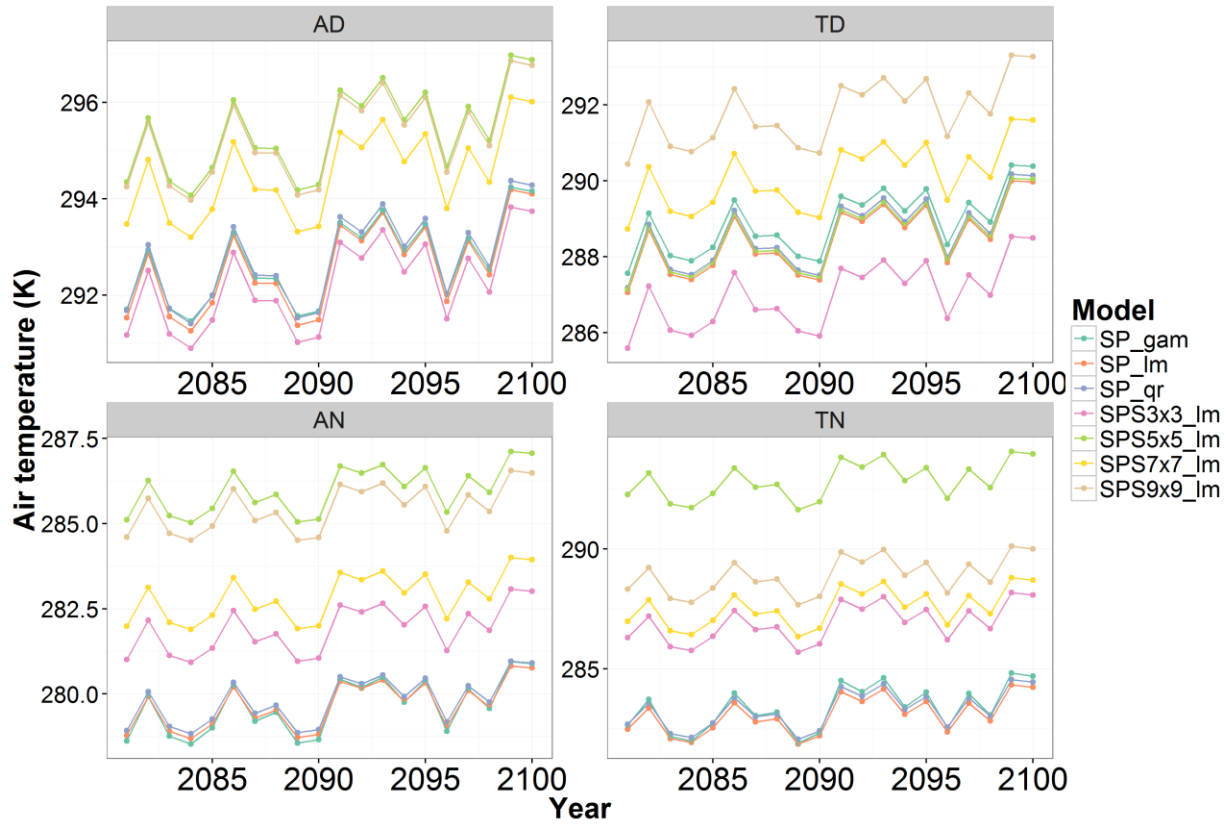
Figure 3.7. Yearly averaged air temperature for the period 2081 to 2100 as predicted by models considered in the direct approach of the application of SP method.
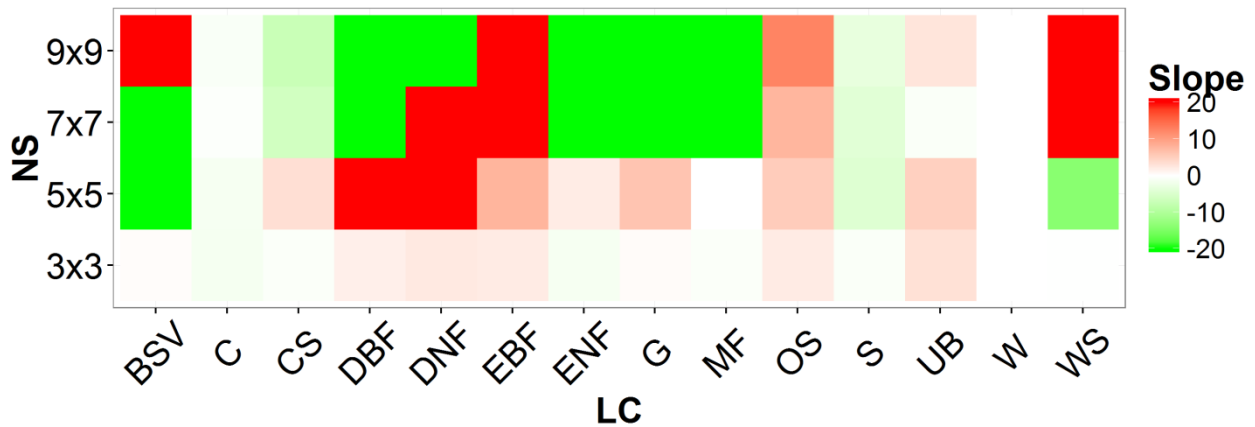
Figure 3.8. Rate of change in temperature with increase in neighborhood land-cover fraction for different land-cover classes (LC) at different neighborhood scales (NS).

Finally the influence of choice of downscaling approach on future projections is analyzed. Annual mean air temperature projected by models 1 and 14 from direct (referred as "dir") and indirect (referred as "ind") approaches are shown in figure 3.9. The mean air temperature projected by the direct approach is found to be 1 K higher than those projected by the indirect approach. This difference in projections is found to occur prominently because direct approach provides projections only for pixels belonging to land-cover classes: C and G because all recording stations are found to be located on one of these two land-cover classes. On the other hand, indirect approach provides projections for all land-cover classes (as seen in Figure C1). This can induce bias in the air temperature projections due to differences in the data distribution of the downscaled outputs obtained from the two approaches. To highlight this bias, projections from indirect approach are calculated neglecting all land-cover classes except C and G. The projections from this experiment are referred as "ind.red" in figure 3.9. It can be noticed that neglecting other land-cover classes increases the mean projected air temperature from indirect approach by 1.1 K and the difference between the mean air temperature projected by the two approaches reduces to 0.3 K. This suggests that the observed differences in projections between direct and indirect approaches are associated with the differences in the distribution of data considered in the two approaches.
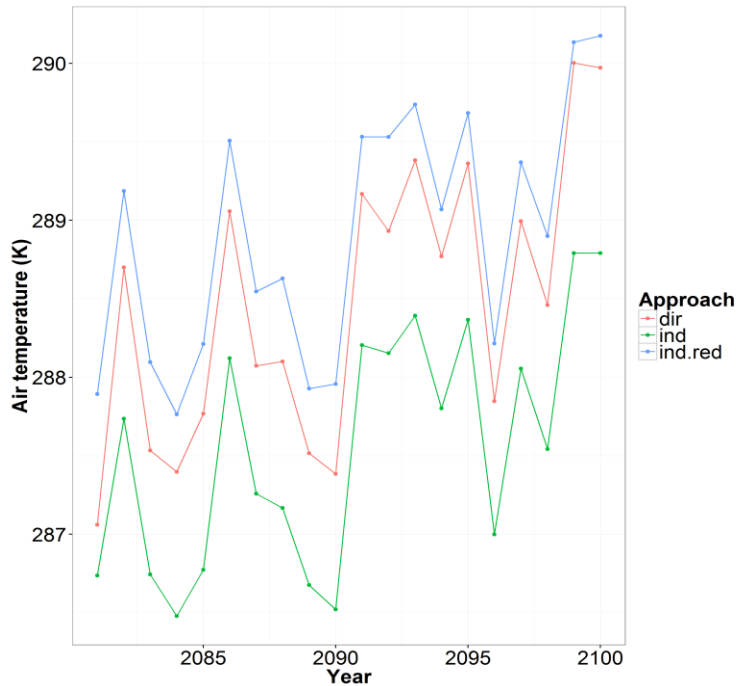
Figure 3.9. Yearly averaged air temperature projections for the period 2081 to 2100 as predicted following the direct approach (dir), indirect approach (ind) and indirect approach considering land-cover classes present in direct approach based projections (ind.red).

3.6 Conclusions

SP method introduced in Gaur and Simonovic (2016) has been used in this study to downscale climate model based air temperature data. The method is found to perform very well in both spatial and temporal robustness tests which test the downscaling efficiency of the models considered in this study. An ensemble of SP model variants are formed by considering two different downscaling approaches, two methods, three functional forms and a set of different predictor variables. Each model of this ensemble is evaluated and compared for their downscaling efficiency. Following key results are obtained:

- Superior model efficiency is obtained from the direct approach than the indirect approach.

- Superior model efficiency is obtained when neighborhood pixel information is taken into account (SPS method) than when it is ignored (SP method).

- Superior model efficiency is obtained when GAMs are used as functional form, followed by LR, followed by QR.

- While performing downscaling using indirect approach, significantly superior model efficiency is obtained when AVs are considered as predictor variables in addition to surface temperature.

The models are thereafter used to downscale future air temperature projections made by a climate model: FGOALS-s2 for an emission scenario: RCP8.5 to evaluate the sensitivity of downscaling approaches, methods and functional forms on future temperature projections. Following results are obtained from the sensitivity analysis:

- Model projections are most sensitive to neighborhood scale considered for analysis, followed by downscaling approaches, followed by functional forms considered.

- High sensitivity of future projections to neighborhood scales can be attributed to significantly different influences of surrounding land-cover classes on the reference pixel at different neighborhood scales.

- Differences between future projections obtained from the direct and indirect approaches can be attributed to the differences in total number of land-cover classes they represent in their projections. Indirect approach provides a more comprehensive picture of changes across the study region as compared to the direct approach. This happens because

remotely sensed data are used to calibrate the models considered in the indirect approach as compared to the direct approach where recording station data is used to calibrate them.

Results from this study support the hypothesis made in Gaur and Simonovic (2016) that local scale climate can be modelled using large scale climate data and land-cover, elevation properties of the location of interest. Further the work also introduces SPS method and highlights the impact of considering neighborhood pixels on projected future temperature. Future work involves the usage of SP and SPS methods towards downscaling climate model based precipitation data as well as to test the applicability of these models at areas outside the region analyzed in this study.

References

Anderson S (2002) An evaluation of spatial interpolation methods on air temperature in Phoenix, AZ. Department of Geography, Arizona State University, Accessed Nov 7 2015. [Available online at: http://www.cobblestoneconcepts.com/ucgis2summer/anderson/anderson.htm].

Argueso D, Evans JP, Fita L, Bormann KJ (2014) Temperature response to future urbanization and climate change. Climate Dynamics 42(7-8): 2183-2199.

Benali A, Carvalho AC, Nunes JP, Carvalhais N, Santos A (2012) Estimating air surface temperature in Portugal using MODIS LST data. Remote Sensing of Environment 124: 108-121.

Chen FW, Liu CW (2012) Estimation of the spatial rainfall distribution using inverse distance weighting (IDW) in the middle of Taiwan. Paddy and Water Environment 10(3): 209-222.

Debele B, Srinivasan R, Yves PJ (2007) Accuracy evaluation of weather data generation and disaggregation methods at finer timescales. Advances in Water Resources 30(5): 1286–1300.

Efe SI, Eyefia AO (2013) Urban Effects on the Precipitation of Benin, Nigeria. American Journal of Climate Change 3: 8-21.

Emamifar A, Rahimikhoob A, Naroozi AA (2013) Daily mean air temperature estimation from MODIS land surface temperature products based on M5 model tree. International Journal of Climatology 33: 3174-3181.

Encyclopedia of Saskatchewan (2015) Climate. University of Regina, Accessed Nov 7 2015. [Available online at http://esask.uregina.ca/entry/climate.html].

Fall S, Niyogi D, Gluhovsky A, and Coauthors (2010) Impacts of land use land cover on temperature trends over the continental United States: assessment using the North American Regional Reanalysis. Int. J. Climatol. 30: 1980–1993.

Fowler, HJ, Blenkinsop S, Tebaldi C (2007) Review: Linking climate change modeling to impacts studies: Recent advances in downscaling techniques for hydrological modeling. Int. J. Climatol. 27: 147–178.

Gaur A, Simonovic SP (2016) A Statistical Model for Physically Representative Downscaling of Surface Temperatures. Manuscript under review for publication in Climate Dynamics.

Goetz SJ, Prince SD, Small JD (2000) Advances in satellite remote sensing of environmental variables for epidemiological applications. Advances in Parasitology 47: 289–307.

Hale RC, Gallo KP, Owen TW, Loveland TR (2006) Land use/land cover change effects on temperature trends at U.S. Climate Normals stations. Geophysical Research Letters 33: L11703.

Hale RC, Gallo KP, Loveland TR (2008) Influences of specific land use/land cover conversions on climatological normals of near-surface temperature. Journal of Geophysical Research 113: D14113.

Hastie and Tibshirani (1990) Generalized Additive Models. Chapman and Hall.

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

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

Kishtawal CM, Niyogi D, Tewari M, et al. (2010) Urbanization signature in the observed heavy rainfall climatology over India. Int. J. Climatol. 30: 1908–1916.

Kusaka H, Hara M, Takane Y (2012) Urban Climate Projection by the WRF Model at 3-km Horizontal Grid Increment: Dynamical Downscaling and Predicting Heat Stress in the 2070's August for Tokyo, Osaka, and Nagoya Metropolises. Journal of the Meteorological Society of Japan 90B: 47-63.

Kusaka H, et al. (2014) Mechanism of Precipitation Increase with Urbanization in Tokyo as Revealed by Ensemble Climate Simulations. J. Appl. Meteor. 53:824–839.

Li W, Chen S, Chen G, et al. (2011) Urbanization signatures in strong versus weak precipitation over the Pearl River Delta metropolitan regions of China. Environmental Research Letters 6: 034020.

Lin Y, Liu A, Ma E, et al. (2013) Impacts of Future Urban Expansion on Regional Climate in the Northeast Megalopolis, USA. Advances in Meteorology 362925.

Lowry WP (1977) Empirical estimation of the urban effects on climate: A problem analysis. J. Appl. Meteor. 16:129–135.

Mesinger F, DiMego G, Kalnay E, et al. (2006) North American Regional Reanalysis. Bull. Amer. Meteor. Soc. 87: 343–360.

Oke TR (1982) The energetic basis of the urban heat island. Quart. J. Roy. Meteor. Soc. 108:1–24.

Qing BAO, Pengfei L, Tianjun Z, et al. (2012) The Flexible Global Ocean-Atmosphere-Land System model Version: FGOALS-s2. AAS CMIP5 Special Issue.

Rao GSP, Jaswal AK, Kumar MS (2004) Effects of urbanization on meteorological parameters. *Mausam* 55: 429-440.

Salathé EP (2003) Comparison of various precipitation downscaling methods for the simulation of streamflow in a rainshadow river basin. International Journal of Climatology 23: 887–901.

Schoof JT (2013) Statistical downscaling in climatology. Geogr. Compass 7: 249–265.

Shepherd JM (2005) A Review of Current Investigations of Urban-Induced Rainfall and Recommendations for the Future. Earth Interactions 9: 1-27.

Stathopoulou M, Cartalis C, Chrysoulakis N (2006) Using midday surface temperature to estimate cooling degree-days from NOAA/AVHRR thermal infrared data: an application for Athens, Greece. Solar Energy 80(4): 414–422.

Stinsen S, Sandholt S, Norgaard A, Fensholt R, Eklundh L (2007) Estimation of diurnal air temperature using MSG SEVIRI data in West Africa. Remote Sensing of Environment 110: 262–274.

Verburg PH, de Nijs TCM, van Eck JR, Visser H, de Jong K (2004) A method to analyse neighbourhood characteristics of land use patterns. Computers, Environment and Urban Systems 28: 667–690.

Vogt J, Viau AA, Paquet F (1997) Mapping Regional Air Temperature Fields Using Satellite Derived Surface Skin Temperatures. International Journal of Climatology 17: 1559–1579.

Wang T, Hamann A, Spittlehouse DL, Murdock TQ (2011) ClimateWNA – high-resolution spatial climate data for western North America. Journal of Applied Meteorology and Climatology 51:16–29.

White-Newsome JL, Brines SJ, Brown DG, et al. (2013) Validating Satellite-Derived Land Surface Temperature with in Situ Measurements: A Public Health Perspective. Environ. Health Perspect. 121: 925–931.

White R, Engelen G (2000) High-resolution integrated modelling of the spatial dynamics of urban and regional systems. Computers, Environment and Urban Systems 24: 383–400.

Wood AW, Leung LR, Sridhar V, Lettenmaier DP (2004) Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. Climatic Change 62: 189–216.

Wood SN (2000) Modelling and Smoothing Parameter Estimation with Multiple Quadratic Penalties. J.R. Statist. Soc. B 62(2):413-428.

Xu Y, Knudby A, Ho HC (2014) Estimating daily maximum air temperature from MODIS in British Columbia, Canada. International Journal of Remote Sensing 35:24: 8108-8121.

Xu Y, Qin Z, Shen Y (2012) Study on the estimation of near-surface air temperature from MODIS data by statistical methods. International Journal of Remote Sensing 33:24: 7629-7643.

Xue Y, Janjic Z, Dudhia J, et al. (2014) A review on regional dynamical downscaling in intraseasonal to seasonal simulation/prediction and major factors that affect downscaling ability. Atmospheric Research 147-148: 68-85.

CHAPTER 4: Application of Physical Scaling towards downscaling climate model precipitation data

4.1 Introduction

General Circulation Models (GCMs) are mathematical representations of the global climate system. They are used to obtain future climate projections across the globe under probable future greenhouse gas emission trajectories. Owing to computational limitations, GCM simulations are performed at grid-sizes that are typically larger than 1° x 1° spatial scale. This spatial scale is much coarser than that required for local or regional scale climate change impact assessment studies. The process of inferring higher spatial resolution climate projections from climate model outputs is referred to as downscaling in climate science literature. Two broad categories of downscaling methodologies have been adopted till date: statistical downscaling and dynamic downscaling while a few studies have combined the two approaches (for instance Svoboda et al. 2012). Statistical downscaling methods link large scale atmospheric variables with locally observed climate data using statistical methods. On the other hand, dynamic downscaling methods simulate higher resolution climate data using boundary conditions simulated by GCMs as inputs into a high resolution mesoscale physically based model (Maraun et al. 2010).

Statistical downscaling methods used in the past can be grouped into four broad categories: 1) scaling methods, 2) regression based approaches, 3) weather pattern based approaches, and 4) weather generators (Schoof 2013). The difference between scaling and regression approaches is that in scaling approach, the value of low resolution climate variable of interest is directly used to infer local scale value of the climate variable of interest. On the other

hand, regression based approaches employ regression methods to link a range of atmospheric variables with local scale climate. For instance, Bias Correction Spatial Downscaling (BCSD) method is a scaling based downscaling method where GCM outputs of the climate variable of interest are first bias corrected and spatially interpolated across the study region. Thereafter by calculating difference between observed climatic data and interpolated GCM data, a spatial anomaly pattern is obtained. This spatial anomaly pattern is kept constant over the historical and future timelines and downscaled GCM projections across both timelines are obtained (Wood et al. 2004). This method has been used in a range of studies to downscale GCM projections (for example Hayhoe et al. 2008; Maurer et al. 2010; Gutmann et al. 2014). Other scaling based downscaling methods such as the 'local' and 'dynamical' scaling approach (Salathe 2003; Widmann et al. 2003; Schmidli et al. 2006) have also been used in the past to downscale GCM data.

Regression based downscaling approaches have been used extensively to downscale GCM data. One very popular downscaling approach is referred to as Statistical DownScaling Method, SDSM (Wilby et al. 2002). In this method, multiple linear regression relationship is developed between a range of low resolution atmospheric variables (for instance geopotential heights, wind speed etc.) and local scale observed climate data. The relationship is thereafter used to estimate local scale downscaled GCM projections. Generalized Linear Modeling framework for downscaling (Fealy and Sweeney 2007) builds a logistic regression model to model local scale precipitation occurrence and a Generalized Additive Model (GAM) to model wet-day precipitation amount from low resolution climate model derived atmospheric variable data. Other regression based downscaling approaches have used quantile regression (Friederichs and Hense 2007), multiway partial least squares regression (Bergant and Kajfez-

Bogataj 2005), canonical correlation (Hertig and Jacobeit 2008), artificial neural networks (Coulibaly et al. 2005), genetic programming (Coulibaly 2004), support vector machines (Tripathi et al. 2006) and relevance vector machines (Ghosh and Mujumdar 2008) to establish relationship between large scale atmospheric variables and locally observed climate data.

A shortcoming of statistical downscaling approaches is that the downscaled products are not physically based. Physical Scaling (SP) downscaling method attempts to overcome this limitation by including large scale climate, physical parameters like elevation and land-cover, as well as physical neighborhood characteristics into the downscaling process. It has been used to downscale GCM based surface temperature (Gaur and Simonovic 2016a) and air temperature data (Gaur and Simonovic 2016b) in the past and has been found to perform better than a state-of-the-art downscaling method: BCSD. In this study SP method based models are evaluated for their ability to downscale North American Regional Reanalysis (NARR) precipitation data. Their performance is compared with two state-of-the-art statistical downscaling methods: SDSM and GLM. The best performing models are thereafter used to downscale future precipitation projections made by three GCMs under two emission scenarios across the study region. The paper is organized as follows. The study region is described in Section 4.2, followed by data used in Section 4.3. This is followed by a description of models and methods used in Section 4.4 followed by a discussion of results in Section 4.5. Finally conclusions from the study are summarized in Section 4.6.

4.2 Study region

The southern Saskatchewan region of Canada is chosen as the study region in this study. The physiographic setting of the study region is shown in Figure 4.1. The red and blue dots represent locations of precipitation gauging stations located within this region. The selected region is representative of Canadian prairies and is characterized by diverse topography and land-cover. Elevation varies between 240 masl to 1389 masl across the study region. Further all land-cover classes identified in the University of Maryland (UMD) classification scheme (summarized in Table 4.1) are found to be present within the study region. An analysis of MODerate-resolution Imaging Spectroradiometer (MODIS) land-cover data (MCD12Q1) between 2006 and 2013 suggests that land-cover classes: Cropland, Evergreen Needle-leaf Forest, Grassland and Mixed Forest constitute approximately 90% of the study region.

Table 4.1. Land-cover classes identified in the UMD classification system. Abbreviations used for different land-cover classes in this study are provided within brackets.

S.No	UMD classes
1	Water (W)
2	Evergreen Needleleaf Forest (ENF)
3	Evergreen Broadleaf Forest (EBF)
4	Deciduous Needleleaf Forest (DNF)
5	Deciduous Broadleaf Forest (DBF)
6	Mixed Forest (MF)
7	Closed Shrublands (CS)
8	Open Shrublands (OS)
9	Woody Savannas (WS)
10	Savannas (S)
11	Grasslands (G)
12	Croplands (C)
13	Urban and Built-up (UB)
14	Barren or Sparsely Vegetated (BSV)

Figure 4.1. Physiographic details of the study region.

The region is land-locked and encompasses many small lakes and rivers. It experiences a continental climate. Large fluctuations in temperature are observed owing to the land-locked location of the region. The region receives almost two-thirds of its precipitation during the summer season, which usually occurs due to large scale convective and cyclonic systems. Significant spatial variability in precipitation is also observed across this region. Snow cover plays a critical role in shaping the hydro-meteorology of the region as this region stays snow-covered almost six months a year (Encyclopedia of Saskatchewan 2016).

The selected region is devoid of any complex physical systems like complex topography, sea coast etc. around it. The motive behind the selection of this region is to evaluate SP method in an isolated and simple region. It is planned that model efficiency will be tested on more climatologically complex regions in future.

4.3 Data used

The following data has been used in this study:

SRTM elevation data: The National Aeronautics and Space Administration (NASA) Shuttle Radar Topographic Mission (SRTM) elevation product (Jarvis et al. 2008) is used in this study. This data has a spatial resolution of 90 m.

MODIS land-cover data: The MODIS land-cover data product (MCD12Q1) in UMD classification scheme is used in this study (LP DAAC 2001). A list of land-cover classes identified in UMD classification scheme is provided in Table 4.1. The data is available in 500 m spatial resolution at annual timesteps. Land-cover data for the period 2006-2012 is selected for analysis. Land-cover for the year 2013 is considered to be the same as that of year 2012 since data for that year is not available from MODIS data repository. This is a reasonable assumption since land-cover in the study region has not changed drastically at annual time-steps in the past (Gaur and Simonovic 2016a).

Gauged daily precipitation data: Daily precipitation data gauged at 57 locations within the study region over the period 2006-2013 is acquired from Environment Canada. The data can be accessed at: http://climate.weather.gc.ca/. The distribution of gauging stations across the study region is shown in Figure 4.1. Using MODIS land-cover data it is found that these gauging stations are associated with UMD land-cover classes: S, OS, G, DNF, UB and C.

NARR climatic and atmospheric data: Daily precipitation rate (P), specific humidity (shum), high cloud area fraction (hcdc), medium cloud area fraction (mcdc), low cloud area fraction (lcdc), air temperature (air) and geopotential height (hgt) data for the period 2006-2013 is acquired from NARR data repository (Mesinger et al. 2006). Data for geopotential height and specific humidity are collected at three vertical levels: 1000 hpa, 850 hpa and 500 hpa while near surface values are extracted for other atmospheric variables. The selection of these large scale atmospheric variables is made keeping in mind the recommendations made in Wilby et al. (2002) as well as the data available in the NARR data repository.

Future land-cover projections: Future land-cover data for the period 2014-2100 for climate models listed in Table 4.2 and emission scenarios: RCP 2.6 and RCP 8.5 are taken from Gaur and Simonovic (2016a). In Gaur and Simonovic (2016a) the 500 m future annual land-cover data is generated by downscaling and reclassifying the future harmonized land-use projections discussed in Hurtt et al. (2011). Future land-cover data is available in UMD classification system (Table 4.1).

Table 4.2. List of GCMs considered for analysis in this study.

GCM	Model	Resolution	Source
1	IAP-FGOALS	$1.66° \times 2.81°$	Institute of Atmospheric Physics, Chinese Academy of Sciences, China
2	MRI-CGCM3	$1.08° \times 2.16°$	Meteorological Research Institute, Japan
3	NorESM1-M	$2° \times 2°$	Norwegian Climate Centre, Norway

Future GCM precipitation data: GCM based daily precipitation data for the period 2006-2100 are collected from the Coupled Model Intercomparison Project-Phase 5 (CMIP5) data repository (Taylor et al. 2012). Data corresponding to climate models listed in Table 4.2 and for two Representative Concentration Pathways (RCPs): RCP 2.6 and RCP 8.5 are acquired.

The choice of climate models is made based on the availability of future land-cover data as developed in Gaur and Simonovic (2016a).

4.4 Models and methods

Three downscaling methods are selected for evaluation in this study. They are described below:

4.4.1 Statistical DownScaling Method (SDSM): Downscaling process by SDSM involves following steps (Wilby et al. 2002):

- Selection of relevant large scale atmospheric predictor variables: This is done by accessing correlation between large scale predictor variables and locally observed climate data. A suitable correlation threshold is chosen to select most relevant atmospheric variables that are later used to model local climate.

- Formulation of regression model: Multiple linear regression relationship between selected large scale atmospheric variables (predictors) and locally observed climate (predictant) data is formulated next. Model is formulated over the chosen calibration period and used to predict local scale climate for the validation period.

- Accounting for internal variability: Multiple realizations of the predicted data are generated using a weather generator in order to account for the internal variability of the climate system.

Since all generated realizations are supposed to have same statistical properties and since the objective of this research is to evaluate SDSM downscaled output, weather generator step is omitted and scaled data is directly used for evaluation. Further several studies have pointed out that the process of initial screening of atmospheric variables is subjective in nature and

96

that this step has significant implications on downscaled outputs (for instance Gagnon et al. 2005), in this study two different versions of SDSM model are evaluated: one with initial screening of atmospheric variables (referred as SDSM.sig hereafter) and one without an initial screening of atmospheric variables (referred as SDSM hereafter).

4.4.2 Generalized Linear Modeling (GLM): Downscaling by GLM involves following steps (Fealy and Sweeney 2007):

- Selection of large scale atmospheric predictor variables: As with SDSM this step involves selecting atmospheric predictor variables that are highly correlated with locally observed climate data.

- Formulation of precipitation occurrence model: A logistic regression approach is employed to simulate wet-dry sequences of precipitation. The formulation can be mathematically expressed as:

$$\ln\left(\frac{P}{1-P}\right) = B_0 + B_1 x_1 + \ldots\ldots B_n x_n \tag{4.1}$$

Where, P denotes the probability of a precipitation event and x denotes independent atmospheric variables selected for analysis. Variable n denotes the number of atmospheric variables selected for prediction.

Formulation of precipitation amounts model: The precipitation amounts model is formulated as a Generalized Additive Model (GAM) between wet day precipitation amount and selected large scale atmospheric variables. The mathematical formulation of the GAM can be expressed as:

$$g(\mu) = f_0 + f_1(x_1) + f_2(x_2) + \ldots f_n(x_n) \tag{4.2}$$

Where g is the link function and f_0, f_1.. f_n represent the non-parametric smooth function associated with n atmospheric variables. In this study, the smooth function is fit using penalized likelihood maximization algorithm. The penalized likelihood maximization algorithm is a variant of maximum likelihood estimation algorithm and applies a tradeoff between model fit wiggliness and goodness of fit by incorporating a penalty function (Wood 2000).

Again in this study two different versions of the GLM model are considered: one with initial screening of atmospheric variables (referred as GLM.sig hereafter) and one without an initial screening of atmospheric variables (referred as GLM hereafter).

4.4.3 SP method based models

SP method utilizes bilinearly interpolated climate model data and physical characteristics of the location of interest as well as its neighborhood to downscale GCM data. Several SP method based models have been explained in Gaur and Simonovic (2016a; 2016b). These have been included in the model ensemble considered for evaluation in this study. Following models have been considered:

SP method based model

Downscaling by SP method is described in details in Gaur and Simonovic (2016a). In this method a multiple linear regression model is formulated with observed climate data as the predictant variable and bilinearly interpolated climate model data, elevation and land-cover as predictor variables. In this study a variant of the SP method with GAMs

as regression function has been used. It has been found in Gaur and Simonovic (2016b) that use of GAM as regression function improves the performance of SP method towards predicting air temperatures. Further since precipitation follows a non-gaussian distribution, it is better to use a regression function which doesn't make gaussian distribution assumption for variables. The downscaling process involves two steps of formulating a precipitation occurrence and amounts model. The steps are similar to GLM method however here model formulation is based on SP method. The mathematical formulation of the precipitation occurrence and amounts model is provided in equations 3 and 4 respectively:

$$\ln\left(\frac{P_{obs}}{1-P_{obs}}\right) = B_0 + B_1 P_{\text{mod}} + B_2 E_p + B_3 LC_p \tag{4.3}$$

$$g(P_{obs,wet}) = B_0 + f_1(P_{\text{mod},wet}) + f_2(E_p) + f_3(LC_p) \tag{4.4}$$

Where, P denotes precipitation, E denotes elevation (masl), LC denotes categorical land-cover variable and B denote regression parameters. Subscript *obs* and *mod* describe if the climatic data is observed or model based respectively. In case of climate model, P_{mod} represents bilinearly interpolated climate model data at a pixel. Subscript p indicates that the data used is a pixel scale data whereas subscript *wet* denotes values on wet days only (i.e. days with more than 0.1 mm of precipitation). Variables g and f represent the link function and smoothing functions respectively.

Two variants of SP method are also considered in this study. First model ignores land-cover as predictor in equations 4.3 and 4.4 and is referred as SP_LC while second

model ignores both land-cover and elevation as predictors in equations 4.3 and 4.4. Latter model is referred to as SP_LC_elev in this paper.

SP method with Surrounding pixel information (SPS) method

The SPS method is a modified version of the SP method in that it incorporates land-cover and elevation properties of the neighborhood pixels in the SP method definition. Gaur and Simonovic (2016b) finds that the inclusion of neighborhood information improves the efficiency of SP method towards downscaling NARR air temperature data by upto 15%. In this study a GAM based version of the SPS model is considered. Again it involves a two-step process of simulating precipitation occurrence and amounts using SPS method. The mathematical formulation of the SPS method can be expressed as:

$$\ln\left(\frac{P_{obs}}{1-P_{obs}}\right) = B_0 + B_1 P_{\mathrm{mod}} + B_2 E_p + B_3 LC_p + B_4 Fr_{W,s} + ... + B_{21} Fr_{BSV,s} + B_{22} R_{E,s} \quad (4.5)$$

$$g(P_{obs,wet}) = B_0 + f_1(P_{\mathrm{mod},wet}) + f_2(E_p) + f_3(LC_p) + f_4(Fr_{W,s}) + ... + f_{21}(Fr_{BSV,s}) + f_{22}(R_{E,s}) \quad (4.6)$$

Where, *g* and *f* denote link and smoothing function respectively. Predictors $Fr_{W,s}$, ...$Fr_{BSV,s}$ represent the fraction of total area surrounding the reference pixel that is occupied by *Water,....Barren and Sparsely Vegetated* land-cover classes respectively. The value of predictors: $Fr_{W,s}$, ...$Fr_{BSV,s}$ is between 0 and 1 and they add up across all neighborhood land-cover classes to give a value of 1. Neighborhood elevation information is incorporated by including a predictor $R_{E,s}$ which represents the ratio between reference pixel elevation and mean elevation of pixels surrounding the reference pixel. The value of each neighborhood predictor is calculated at a certain neighborhood scale (denoted by *s* in equation 4.6). In this study four neighborhood

scales: 3x3, 5x5, 7x7 and 9x9 are considered. These neighborhood scales have been used in previous studies (Verberg et al. 2004; Gaur and Simonovic 2016b) and have been adopted in this study as well. Configuration of neighborhood scales considered in this analysis is shown in Figure 4.2. The reference pixel is shown in red. Neighborhood pixels encompassed in 3x3, 5x5, 7x7 and 9x9 neighborhood scale are shown in yellow, green, orange and blue respectively. Areas encompassed in higher neighborhood scales are inclusive of smaller neighborhood scales. This means that neighborhood area of 5x5 scale encompasses the area associated with 3x3 neighborhood scale plus the yellow area. Models calibrated at 3x3, 5x5, 7x7 and 9x9 neighborhood scales are referred as SPS3x3, SPS5x5, SPS7x7 and SPS9x9 respectively in this paper.

A list of all models that are evaluated in this study is provided in Table 4.3.
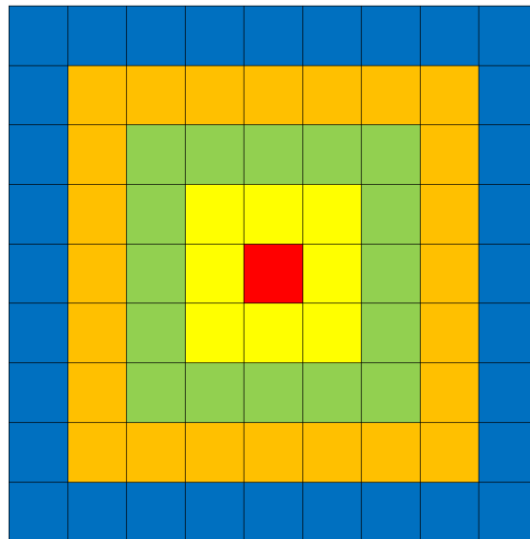
Figure 4.2. Neighborhood scales considered for analysis in this study. Reference pixel is shown in red color while pixels encompassed in 3x3, 5x5, 7x7 and 9x9 neighborhood scales are shown

in yellow, green, orange and blue respectively. Areas encompassed by higher neighborhood

scales are inclusive of the smaller neighborhood scales.

Table 4.3. Downscaling models evaluated in this study.

S.No	Model name	Model name (short)	Predictors
1	SP	M1	P, LC, E
2	SP_LC	M2	P, E
3	SP_LC_elev	M3	P
4	SPS3x3	M4	P, LC, E, NLC_{3x3}, NE_{3x3}
5	SPS5x5	M5	P, LC, E, NLC_{5x5}, NE_{5x5}
6	SPS7x7	M6	P, LC, E, NLC_{7x7}, NE_{7x7}
7	SPS9x9	M7	P, LC, E, NLC_{9x9}, NE_{9x9}
8	SDSM	M8	wnd, rhum, prmsl, lcdc, shum1000hpa, mcdc, hcdc, air, shum850hpa, shum500hpa, hgt1000hpa, hgt850hpa, hgt500hpa
9	GLM	M9	
10	SDSM.sig	M10	lcdc, mcdc, hcdc
11	GLM.sig	M11	

4.5 Results and discussion

The models described before are formulated for snow-free (referred as *sf*) and snow-covered

(referred as *sc*) months separately. In this study May to September are considered as snow-

free months while October to April are considered as snow-covered months. Since SP

method and GLM based models both employ GAM as the regression function, the same is

used to build relationship between low resolution atmospheric variables and locally observed

climatic data in the SDSM model. This is done to maintain regression function consistency

among all models being evaluated in this study so that an unbiased evaluation of downscaling

methodologies can be made.

The choice of predictors for SDSM.sig and GLM.sig models is made by analyzing monthly correlation between atmospheric variables and locally observed precipitation data at all gauging stations located within the study region. Results are presented in Figure 4.3 where average spearman correlation coefficients between atmospheric variables and local precipitation are plotted for all months. Highest correlation values are obtained in the case of cloud-cover variables: high cloud area fraction (hcdc), medium cloud area fraction (mcdc) and low cloud area fraction (lcdc). Therefore they are selected as atmospheric predictor variables for performing downscaling by SDSM.sig and GLM.sig models. For SDSM and GLM models, all atmospheric variables listed in Table 4.3 are considered.
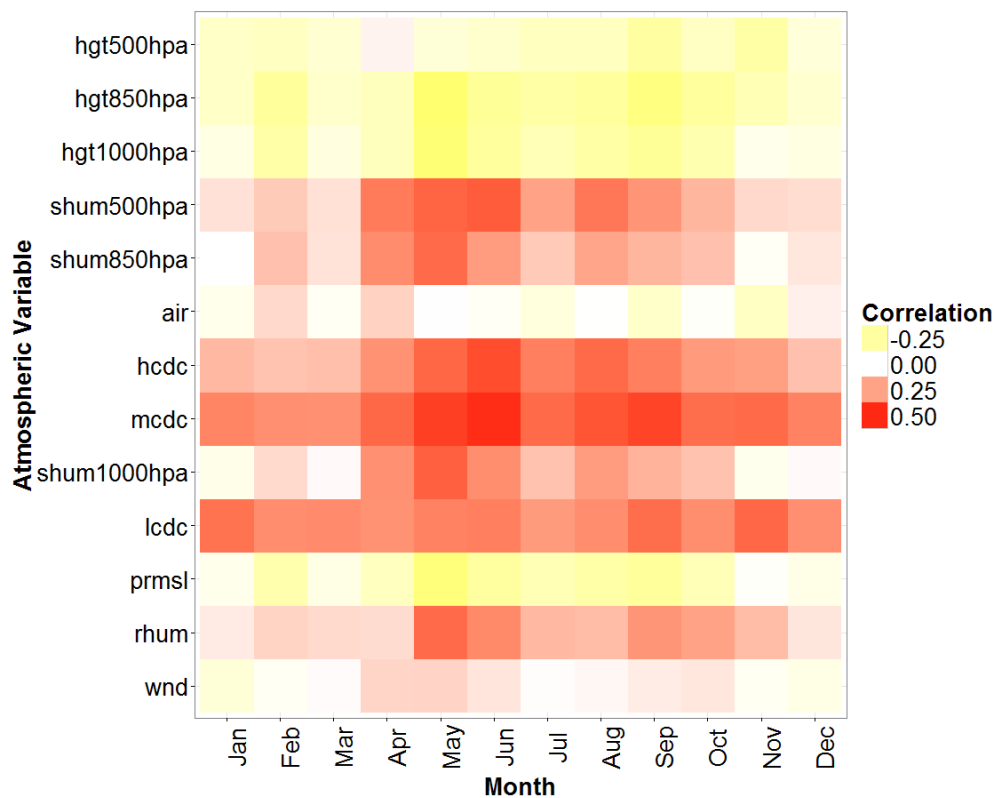
Figure 4.3. Monthly correlation between low resolution atmospheric variable data and locally observed precipitation.

Two different tests of robustness are performed: 1) Test for temporal robustness (TR) and 2) Test for spatial robustness (SR). In the temporal robustness test downscaling models are calibrated over the period: 2006-2010 and validated over the period 2011-2013. On the other hand, in the spatial robustness test the downscaling models are calibrated across 29 (out of 57) stations located across the study region and validated across the rest of the gauging stations. The downscaled precipitation are evaluated on the basis of the downscaling model's ability to simulate following seven precipitation based indices: 1) Spearman correlation coefficient between model simulated and observed data (sp.cor), 2) fraction of dry days i.e. fraction of days with less than 0.1 mm of rainfall (ddays), 3) maximum precipitation intensity (ppt.max), 4) mean wet day precipitation (ppt.wet), 5) total number of one-day precipitation events (p1d), 6) total number of 2-4 day precipitation events (p2to4d) and 7) total number of 5 or more day precipitation events (p5d).

While calibrating SP method and GLM based models, probability predictions made in the occurrence model are associated with an occurrence (1) or no-occurrence (0) value using a threshold value such that:

$$f(p) = \begin{cases} 1 \text{ if } p >= p_{threshold} \\ 0 \text{ if } p < p_{threshold} \end{cases} \tag{4.7}$$

Where p denotes predicted probabilities as obtained from the occurrence model and $p_{threshold}$ denotes the threshold probability value chosen for analysis. In this study a series of $p_{threshold}$ values ranging between 0 and 1 are tested to select a threshold probability value that provides maximum prediction accuracy to the SP method and GLM based models.

The variation of model efficiencies with probability threshold values for SP method based models and GLM models for both snow-cover states and robustness tests is presented in Figure 4.4. Twenty-one probability threshold values evenly spaced between 0 and 1 at a spacing of 0.05 are considered for analysis. Model efficiency is calculated by evaluating the percentage of total data length correctly predicted by the calibrated model on validation time-series. From the plots, it can be seen that occurrence model performance for SP and GLM based models vary significantly with the choice of probability threshold value. Further minor variations in model efficiency are also observed with differences in snow-cover state, robustness test and downscaling model considered. Optimal threshold value for each model, snow-cover state and robustness test combination is used for making prediction from these models. A summary of these optimal threshold values is presented in Table 4.4. It is noticed that threshold values for GLM models are higher than the threshold values of SP method based models.
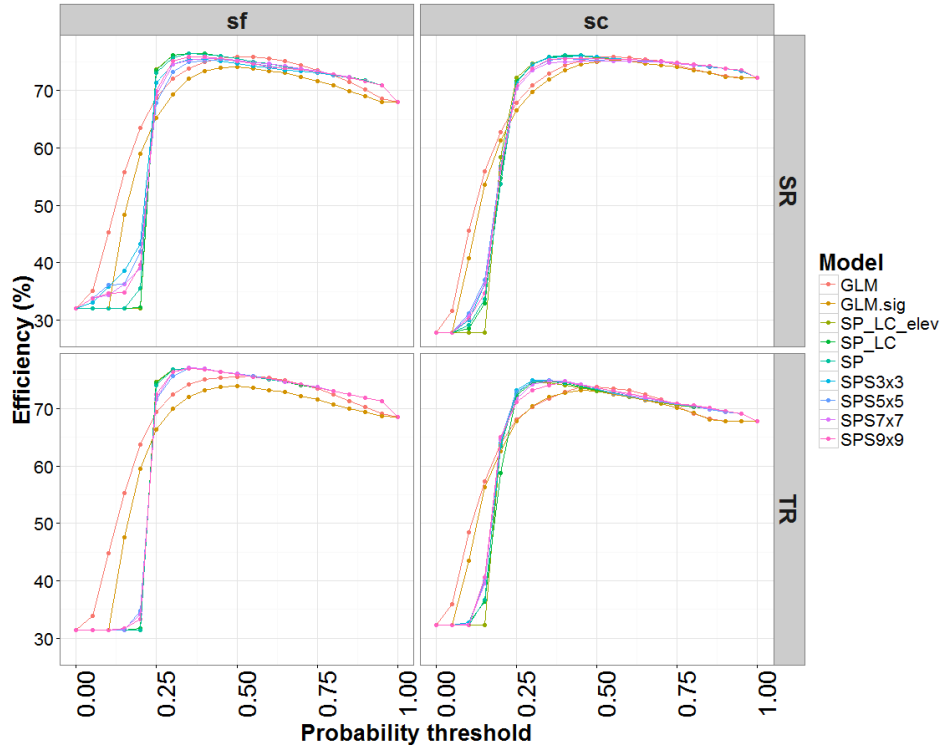
Figure 4.4. Variation of model efficiencies with probability threshold values for GLM and SP method based models. Efficiency values are presented for snow-free (sf) and snow-covered (sc) timelines for spatial robustness (SR) and temporal robustness (TR) tests.

Table 4.4. Optimum probability threshold values for different models for snow-free (sf) and snow-covered (sc) months, and for TR (SR) tests.

Model	Snow	
	sf	*sc*
GLM	0.55 (0.5)	0.45 (0.55)
GLM.sig	0.5 (0.5)	0.45(0.55)
SP	0.35 (0.35)	0.35 (0.4)
SP-LC	0.35 (0.35)	0.35 (0.4)
SP-LC-elev	0.35 (0.35)	0.3 (0.45)
SPS3x3	0.35 (0.4)	0.35 (0.45)
SPS5x5	0.4 (0.45)	0.35 (0.4)
SPS7x7	0.35 (0.45)	0.35 (0.55)
SPS9x9	0.35 (0.35)	0.4 (0.5)

The calibrated models are used to downscale NARR precipitation grid data across the validation period (for TR test) and validation stations (for SR test). The performance of models in terms of rank correlation between modelled and observed data is shown in Table 4.5 for all experiments and snow-cover states. Correlation values averaged across all experiments and snow-cover states are also shown. Best and second best performing models in terms of average performance are highlighted in orange and green respectively. It can be seen that SP method based models majorly outperform both SDSM and GLM models in terms of correlation. The performance of SP models is better in the snow-free months ($rho_{avg.}$ = 0.5) as compared to snow-covered months ($rho_{avg.}$ = 0.4) and better in the TR test ($rho_{avg.}$ = 0.5) than the SR test ($rho_{avg.}$ = 0.4). Following SP method based models, SDSM model is found to perform best, followed by GLM model. Further the performance of SDSM.sig and GLM.sig models are found to be inferior than the SDSM and GLM models.

Table 4.5. Spearman correlation coefficient between model simulated and observed precipitation for models considered in this study. Best and second best model based on average correlation coefficient are highlighted in orange and green respectively. Models M1 to M11 can be referred to from Table 4.3.

Exp	Snow	Models										
		M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11
SR	sf	0.47	0.48	0.48	0.44	0.44	0.44	0.47	0.47	0.46	0.42	0.40
SR	sc	0.38	0.37	0.36	0.36	0.37	0.32	0.35	0.41	0.35	0.38	0.33
TR	sf	0.48	0.48	0.48	0.48	0.48	0.49	0.48	0.44	0.43	0.40	0.37
TR	sc	0.42	0.43	0.44	0.42	0.42	0.42	0.41	0.40	0.40	0.42	0.36
Average		0.44	0.44	0.44	0.43	0.43	0.42	0.43	0.43	0.41	0.41	0.36

SP models perform best in simulating the fraction of total number of dry days in the validation time-series as evident from Table 4.6. The performance of SP models is again found to be better in snow-free months ($RMSE_{avg.}= 0.1$) than the snow-covered months ($RMSE_{avg.}= 0.5$) and in the TR test ($RMSE_{avg.} = 0.11$) than in the SR test ($RMSE_{avg.}= 0.12$). Among other models, GLM model is also found to perform well ($RMSE_{avg.} = 0.11$) followed by SDSM ($RMSE_{avg.} = 0.5$). Again the performance of SDSM.sig and GLM.sig models towards simulating dry day fraction is found to be inferior than the SDSM and GLM models.

Table 4.6. Dry day fraction as obtained from observed data and as well as downscaled precipitation obtained using models considered for analysis. Best and second best model based on average bias are highlighted in orange and green respectively. Models M1 to M11 can be referred to from Table 4.3.

Exp	Snow	obs	Models										
			M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11
SR	sf	0.68	0.76	0.76	0.76	0.80	0.81	0.81	0.76	0.21	0.76	0.11	0.79
SR	sc	0.72	0.85	0.85	0.87	0.87	0.83	0.89	0.88	0.25	0.86	0.20	0.85
TR	sf	0.69	0.77	0.77	0.77	0.77	0.80	0.77	0.77	0.18	0.81	0.10	0.81
TR	sc	0.68	0.80	0.81	0.76	0.80	0.80	0.79	0.82	0.24	0.77	0.17	0.84
Average RMSE			0.11	0.11	0.11	0.12	0.12	0.13	0.12	0.47	0.11	0.55	0.13

Maximum precipitation intensity is not simulated satisfactorily by all three types of models considered in this study. This can be seen from Table 4.7 where biases associated with maximum precipitation values are presented. SP model is found to perform best followed by SP_LC_elev model. Among the three types of models, SP method based models are found to perform best, followed by GLM and followed by SDSM model. SP model performance is found to be significantly better in snow-free months ($RMSE_{avg.} = 54$ mm) than the snow-

covered months (RMSE$_{avg.}$ = 90 mm) and in the TR test (RMSE$_{avg.}$ = 70 mm) than SR test (RMSE$_{avg.}$ = 78 mm). Further the performance of SDSM.sig and GLM.sig models towards simulating maximum precipitation intensity is found to be inferior than SDSM and GLM models.

Table 4.7. Maximum precipitation as obtained from observed data and as well as downscaled precipitation obtained using models considered for analysis. Best and second best model based on average bias are highlighted in orange and green respectively. Models M1 to M11 can be referred to from Table 4.3.

Exp	Snow	obs	Models										
			M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11
SR	sf	101	37.4	36.6	37.2	37.0	36.4	35.9	36.6	14.0	30.0	11.0	13.1
SR	sc	102	13.8	12.4	12.4	12.9	12.4	12.0	11.9	8.8	12.8	6.2	8.4
TR	sf	72	32.1	31.9	31.8	32.0	31.9	31.8	31.8	17.9	28.8	11.3	13.0
TR	sc	102	11.3	10.5	10.5	11.1	11.2	10.7	10.6	7.1	9.4	6.2	7.8
Average RMSE			73.3	74.2	74.1	73.7	74.0	74.4	74.3	83.8	76.3	86.6	84.7

Wet day mean precipitation is simulated reasonably well by SP method based models (Table 4.8). Among the three types of models considered, GLM model is found to perform best (RMSE$_{avg.}$ = 0.8 mm), followed by SP method (RMSE$_{avg.}$ = 1 mm), and followed by SDSM (RMSE$_{avg.}$ = 2.8 mm). In the case of wet day mean precipitation, SP model performance is found to be better in the snow-covered months (RMSE$_{avg.}$ = 0.9 mm) than the snow-free months (RMSE$_{avg.}$ = 1 mm) and better in the TR test (bias$_{avg.}$ = 0.8 mm) than the SR test (bias$_{avg.}$ = 1 mm). Again the performance of SDSM.sig and GLM.sig models towards simulating mean wet day precipitation intensity is found to be inferior than SDSM and GLM models.

Table 4.8. Mean wet-day precipitation as obtained from observed data and as well as downscaled precipitation obtained using models considered for analysis. Best and second best model based on average bias are highlighted in orange and green respectively. Models M1 to M11 can be referred to from Table 4.3.

Exp	Snow	obs	Models										
			M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11
SR	sf	5.82	6.80	6.84	6.81	7.09	7.11	7.17	6.54	2.40	6.83	2.02	6.89
SR	sc	2.85	3.84	3.85	4.11	3.95	3.34	3.84	3.87	1.00	3.43	0.93	3.36
TR	sf	6.02	6.84	6.84	6.84	6.87	7.26	6.79	6.75	2.34	7.15	1.97	7.28
TR	sc	2.87	3.56	3.79	3.47	3.69	3.67	3.60	3.75	1.01	3.17	0.94	3.70
Average RMSE			0.88	0.94	0.95	1.02	1.01	0.99	0.85	2.84	0.82	3.10	0.96

The occurrences of 1-day and 2-4 day precipitation events are best simulated by SP method based models whereas GLM model performs best in simulating 5 or more day precipitation events as evident in Tables 4.9 to 4.11. SDSM model is found to underestimate the occurrence frequency of one-day and 2-4 day precipitation events and overestimate the occurrence frequency of 5 or more day precipitation events. The performance of SP models is again found to be superior in the TR test (RMSE$_{avg.}$ = 154, 420, 132 respectively for 1-day, 2-4 days and 5 or more days precipitation event) than in the SR test (RMSE$_{avg.}$ = 298, 743, 223 respectively for 1-day, 2-4 days and 5 or more days precipitation event). Further SP method based models are found to perform better in snow-free months than in the snow-covered months for 1 day and 2-4 day precipitation events (RMSE$_{avg.}$ found lower by 32 and 178 for 1 day and 2-4 day precipitation events respectively). However they are found to perform better in the snow-covered months than snow-free months for more than 5 day precipitation events (RMSE$_{avg.}$ found lower by 58). The performance of SDSM.sig and GLM.sig models towards simulating 1 day and 2-4 day precipitation events is found to be

inferior than SDSM and GLM models. However in the case of 5 or more day precipitation

events, SDSM.sig model is found to perform better than SDSM model. GLM.sig model still

performs inferiorly to GLM model in the case of 5 or more day precipitation events.

Table 4.9. Total 1-day precipitation events as obtained from observed data and as well as

downscaled precipitation obtained using models considered for analysis. Best and second best

model based on average bias are highlighted in orange and green respectively. Models M1 to

M11 can be referred to from Table 4.3.

Exp	Snow	obs	Models										
			M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11
SR	sf	2907	2893	2896	2905	2573	2483	2526	2771	772	2760	0	2873
SR	sc	2025	1942	1933	1750	1738	1850	1422	1558	488	1987	331	1809
TR	sf	2004	1846	1852	1858	1830	1725	1836	1856	404	1780	85	1910
TR	sc	1431	1382	1433	1597	1346	1345	1316	1211	331	1435	146	1229
Average RMSE			93	89	176	241	272	371	277	1635	135	2040	156

Table 4.10. Total 2-4 day precipitation events as obtained from observed data and as well as

downscaled precipitation obtained using models considered for analysis. Best and second best

model based on average bias are highlighted in orange and green respectively. Models M1 to

M11 can be referred to from Table 4.3.

Exp	Snow	obs	Models										
			M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11
SR	sf	2374	1998	1989	1998	1575	1453	1457	1967	1530	1996	0	1674
SR	sc	1463	787	767	608	608	790	439	513	830	671	639	852
TR	sf	1568	1273	1266	1267	1321	1112	1300	1285	783	973	113	1000
TR	sc	1197	674	724	955	656	652	671	594	518	930	342	578
Average RMSE			490	487	505	656	672	748	615	740	547	1514	626

Table 4.11. Total 5 or more day precipitation events as obtained from observed data and as well as downscaled precipitation obtained using models considered for analysis. Best and second best model based on average bias are highlighted in orange and green respectively. Models M1 to M11 can be referred to from Table 4.3.

Exp	Snow	obs	Models										
			M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11
SR	sf	355	122	115	114	61	64	46	124	1706	164	1	91
SR	sc	198	33	30	11	22	51	13	20	1406	18	1479	21
TR	sf	196	61	64	63	68	38	63	67	1123	50	466	24
TR	sc	161	26	29	46	29	31	29	44	956	66	979	21
Average RMSE			172	174	176	194	192	203	170	1093	157	792	194

Overall SP method based models are found to perform better than the SDSM and GLM based models. This is evident in Figure 4.5 where index specific bias associated with each individual model is normalized and presented. Index specific RMSE values are normalized so that inter-model comparisons can be made taking into consideration all seven indices. A lighter shade represents a better performing model. Overall, based on average normalized RMSE ($RMSE_{ANB}$), models ranked as: 1) SP_LC_elev ($RMSE_{ANB} = 0.02$), 2) SP ($RMSE_{ANB} = 0.03$), 3) SP_LC ($RMSE_{ANB} = 0.04$), 4) SPS9x9 ($RMSE_{ANB} = 0.16$), 5) SPS3x3 ($RMSE_{ANB} = 0.19$), 6) SPS5x5 ($RMSE_{ANB} = 0.22$), 7) GLM ($RMSE_{ANB} = 0.22$), 8) SPS7x7 ($RMSE_{ANB} = 0.28$) and 9) SDSM ($RMSE_{ANB} = 0.91$) in terms of model performance across indices. It can be noticed that SP method based models: SP, SP_LC and SP_LC_elev perform significantly better than all other models considered in this analysis and hence are chosen for making future precipitation projections across the study region.
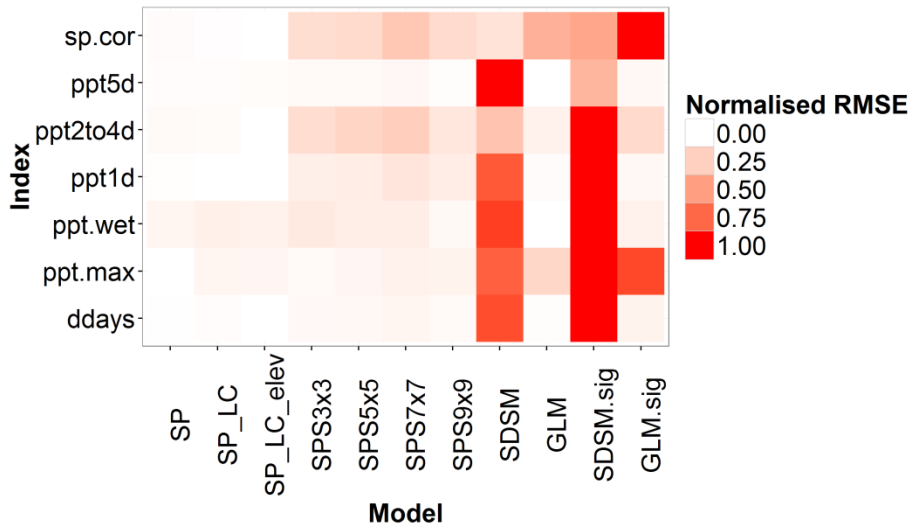
Figure 4.5. Normalized bias associated with different models for indices considered in this study.

Future precipitation projections made by selected GCMs are used to obtain downscaled precipitation projections across the study region. Future land-cover projections are used while making future precipitation projections using SP model. The precipitation gauging stations located within the study region are found to be associated with S, OS, G, DNF, UB and C land-cover classes therefore future predictions by SP model could only be made at pixels belonging to these land-cover classes. In order to maintain consistency, future projections from all three models selected for making future projections are only made at pixels belonging to above mentioned land-cover classes.

Future projections of precipitation in terms of indices: ddays, ppt.max, ppt.wet, ppt1d, ppt2to4d and ppt5d corresponding to each climate model and emission scenario combination are presented in Figure 4.6. Projections corresponding to all GCMs, RCPs, snow-cover

states, land-cover classes and three downscaling models are averaged and discussed hereafter. It is found that between 2014 and 2100, the mean wet day precipitation across the study region is projected to increase at an average rate of 0.002 mm/year ($p < 0.001$) while the dry day fraction is projected to decrease at a rate of 0.0003/year ($p < 0.001$). Maximum precipitation is projected to increase at a rate of 0.008 mm/year ($p = 0.03$). Further the frequency of occurrence of 1 day, 2-4 days and more than 5 day precipitation events is also projected to increase in the future at rates of 0.06/year ($p = 0.001$), 0.06/year ($p = 0.001$) and 0.007/year ($p = 0.4$) respectively.
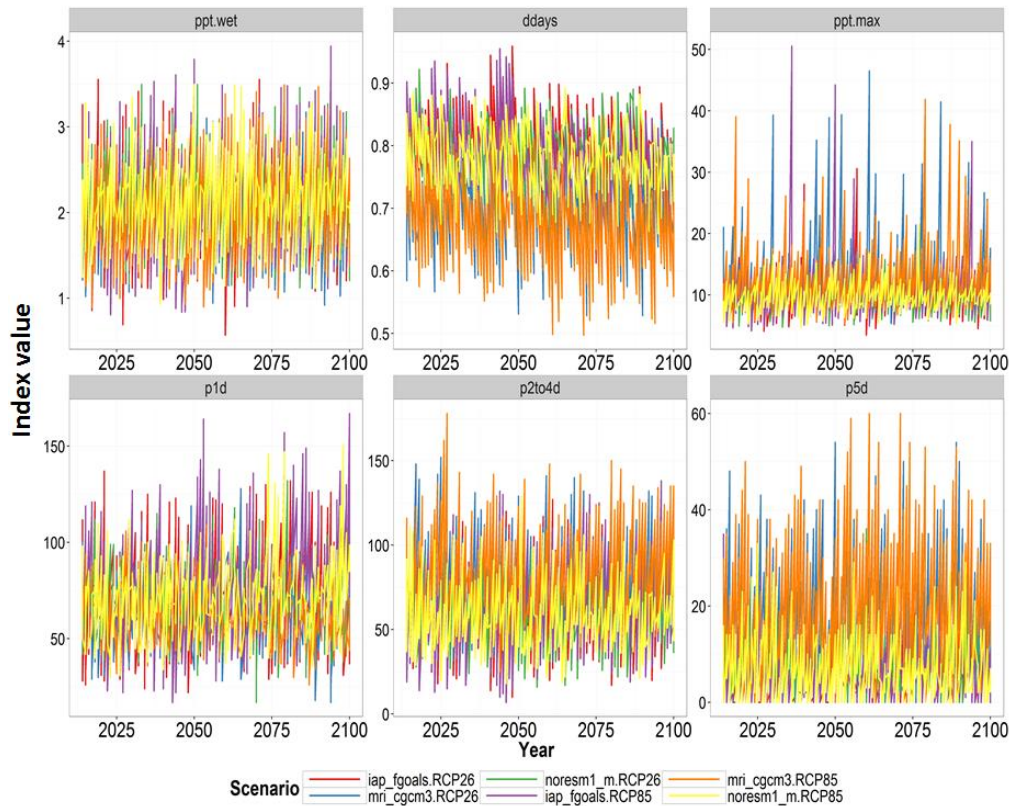
Figure 4.6. Yearly dry day fraction (ddays), maximum precipitation (ppt.max), mean wet day precipitation (ppt.wet), total number of 1 day precipitation events (ppt1d), total number of 2 to 4 day precipitation events (ppt2to4d) and total number of more than 5 day precipitation events

(ppt5d) as projected across the study region. Projections corresponding to all GCMs, RCPs, snow-cover states, land-cover classes and downscaling methods are combined here.

The uncertainty in future precipitation projections as induced by different GCMs, RCPs, snow-cover state (snow-covered or snow-free) and land-cover classes considered in this study are analyzed. Only SP method is considered while performing uncertainty analysis since other methods (SP_LC and SP_LC_elev) don't take land-cover into consideration while downscaling precipitation. The results are presented in Figure 4.7 where variations in different indices as induced by GCMs, RCPs, snow-cover state and land-cover classes are presented. It can be seen that the contribution of different sources towards total uncertainty varies for different indices considered. Snow-cover state is found to be the most important source of uncertainty for three indices: ppt.max, p1d and p2to4d. Land-cover is found to be the most important factor governing mean wet day precipitation. GCMs are found to be the most important source of uncertainty in terms of dry day fraction and more than 5 day precipitation events. Overall emission scenarios are found to be the most insignificant contributor towards total uncertainty while snow-cover state is found to be the most important contributor.
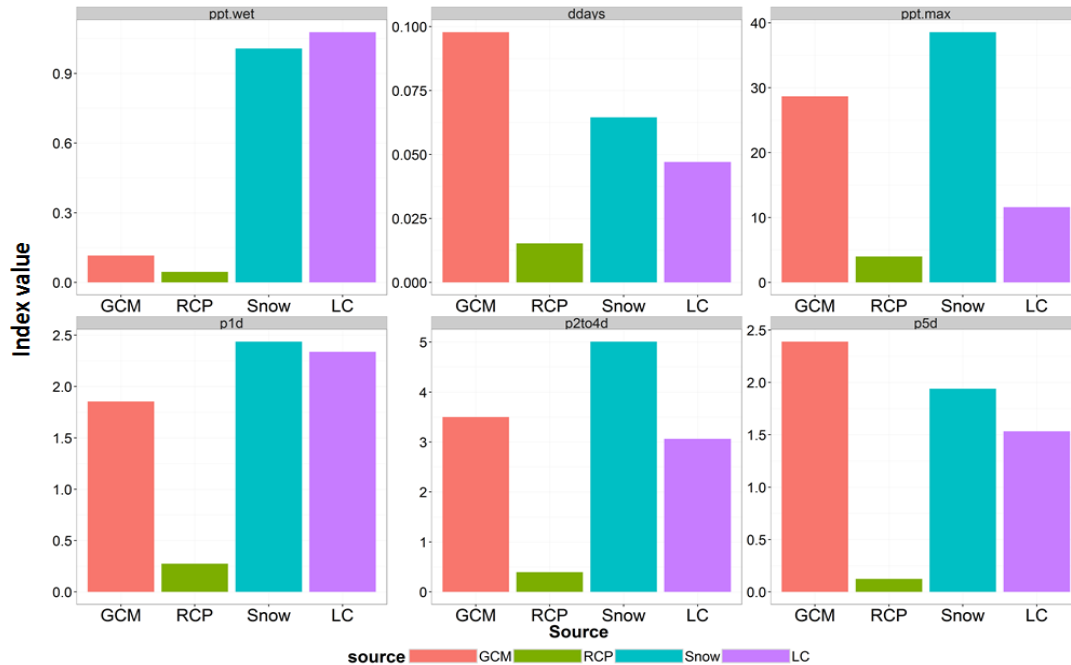
Figure 4.7. Relative contribution of different sources towards uncertainty in future precipitation projections.

4.6 Conclusions

In this study the applicability of SP method (and its extensions) towards downscaling climate model based precipitation data is accessed. The performance of SP method based models is compared with two state-of-the-art statistical downscaling models: SDSM and GLM. Further two variants of SDSM and GLM models are used: one incorporating initial atmospheric variable selection step (SDSM.sig and GLM.sig) and one considering all atmospheric variables for prediction (SDSM and GLM). Models are accessed based on their ability to model seven precipitation based indices. Further model performance in terms of spatial and temporal robustness is accessed. Major conclusions include:

- Overall it is found that most SP method based models outperform SDSM and GLM based models. Further GLM model is found to perform better than the SDSM model.

- It is found that models SP, SP_LC and SP_LC_elev are found to perform significantly better than other models analyzed.

- Inclusion of neighborhood land-cover and elevation characteristics doesn't improve SP method performance.

- The performance of SDSM and GLM models is found to be significantly better than the SDSM.sig and GLM.sig models. This suggests that atmospheric predictor variable step doesn't improve model performance. These results highlight the benefit of considering climate variable of interest as a predictor variable while downscaling GCM outputs.

Best performing downscaling models: SP, SP_LC and SP_LC_elev are thereafter used to downscale future projections made by three GCMs under two extreme emission scenarios: RCP 2.6 and RCP 8.5. Future precipitation projections indicate an increase in precipitation intensity across the study region. Increase in mean and extreme precipitation intensity is projected. Further increase in the frequency of short (1-day), moderately long (2-4 days) and long (more than 5 day) precipitation events is projected. Contribution of GCMs, RCPs, snow-cover state and land-cover classes towards total uncertainty is assessed. It is found that the relative contribution of different sources of uncertainty varies for different precipitation indices considered. Overall snow-cover state of the location of interest is identified as the most important source of uncertainty, followed by GCMs, followed by land-cover classes, followed by the emission scenario.

The results from this study support the findings of Gaur and Simonovic (2016a; b) that SP method can be used to downscale GCM data. Further exploration is required about finding an

appropriate spatial and temporal scale at which model calibration should be performed. This will be the future direction of this work.

References

Bergant K, Kajfez-Bogataj L (2005) N-PLS regression as empirical downscaling tool in climate change studies. Theoretical and Applied Climatology 81: 11–23.

Coulibaly P, Dibike YB, Anctil F (2005) Downscaling precipitation and temperature with temporal neural networks. Journal of Hydrometeorology 6: 483–496.

Coulibaly P (2004) Downscaling daily extreme temperatures with genetic programming. Geophysical Research Letters 31: L16203.

Encyclopedia of Saskatchewan (2016) Climate. University of Regina, Accessed March 7 2016. [Available online at http://esask.uregina.ca/entry/climate.html].

Fealy R, Sweeney J (2007) Statistical downscaling of precipitation for a selection of sites in Ireland employing a generalized linear modeling approach. International Journal of Climatology 27: 2083–2094.

Friederichs P, Hense A (2007) Statistical downscaling of extreme precipitation events using censored quantile regression. Monthly Weather Review 135: 2365–2378.

Gaur A, Simonovic SP (2016a) A Scaling Method for Physically Representative Downscaling of Climate Model Data. Under review in Climate Dynamics.

Gaur A, Simonovic SP (2016b) Extension of SP method and its application towards downscaling climate model based near surface air temperature. Under review in Journal of Applied Meteorology and Climatology.

Ghosh S, Mujumdar PP (2008). Statistical downscaling of GCM simulations to streamflow using relevance vector machine. Advances in Water Resources 31: 132–146.

Gutmann E, Pruitt T, Clark MP, Brekke L, Arnold JR, Raff DA, Rasmussen RM (2014) An intercomparison of statistical downscaling methods used for water resource assessments in the United States. Water Resour. Res. 50: 7167–7186.

Hayhoe K, et al. (2008) Regional climate change projections for the Northeast USA. Mitigation and Adaptation Strategies for Global Change 13 (5): 425-436.

Hertig E, Jacobeit J (2008) Downscaling future climate change: temperature scenarios for the Mediterranean area. Global and Planetary Change 63: 127–131.

Hurtt GC, et al. (2011) Harmonization of land-use scenarios for the period 1500–2100: 600 years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands. Clim. Change 109(1-2): 117–161.

Jarvis A, Reuter HI, Nelson A, Guevara E (2008) Hole-filled SRTM for the globe Version 4, available from the CGIAR-CSI SRTM 90m Database (http://srtm.csi.cgiar.org).

Land Processes Distributed Active Archive Center (LP DAAC) (2001) Land Cover Type Yearly L3 Global 500 m SIN Grid. Version 051. NASA EOSDIS Land Processes DAAC, USGS Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota

(https://lpdaac.usgs.gov), accessed March 14, 2016, at
http://dx.doi.org/10.5067/ASTER/AST_L1B.003.

Maurer EP, Hidalgo GP, Das T, Dettinger MD, Cayan DR (2010) The utility of daily large-scale
climate data in the assessment of climate change impacts on daily streamflow in California.
Hydrol. Earth Syst. Sci. 14: 1125–1138.

Maraun D, et al. (2010) Precipitation downscaling under climate change: Recent developments
to bridge the gap between dynamical models and the end user. Rev. Geophys. 48: RG3003.

Mesinger F, DiMego G, Kalnay E, et al. (2006) North American Regional Reanalysis. Bull.
Amer. Meteor. Soc. 87: 343–360.

Salathé EP (2003) Comparison of various precipitation downscaling methods for the simulation
of streamflow in a rainshadow river basin. International Journal of Climatology 23: 887–901.

Schoof JT (2013) Statistical downscaling in climatology. Geogr. Compass 7: 249–265.

Schmidli J, Frei C, Vidale PL (2006) Downscaling from GCM predictors: a benchmark for
dynamical and statistical downscaling methods. International Journal of Climatology 26:
679–689.

Taylor KE, Stouffer RJ, Meehl GA (2012) An overview of CMIP5 and the experimental design.
Bulletin of the American Meteorological Society 93: 485–498.

Tripathi S, Srinivas VV, Nanjundiah RS (2006) Downscaling of precipitation for climate change
scenarios: a support vector machine approach. Journal of Hydrology 330: 621–640.

Verburg PH, de Nijs TCM, van Eck JR, Visser H, de Jong K (2004) A method to analyse neighbourhood characteristics of land use patterns. Computers, Environment and Urban Systems 28: 667–690.

Widmann M, Bretherton CS, Salathé EP (2003) Statistical precipitation downscaling over the Northwestern United States using numerically simulated precipitation as a predictor. Journal of Climate 16: 799–816.

Wilby RL, Dawson CW, Barrow EM (2002) SDSM – a decision support tool for the assessment of regional climate change impacts. Environmental Modelling & Software 17(2): 145–157.

Wood AW, Leung LR, Sridhar V, Lettenmaier DP (2004) Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. Climatic Change 62: 189–216.

Wood SN (2000) Modelling and Smoothing Parameter Estimation with Multiple Quadratic Penalties. J.R. Statist. Soc. B 62(2):413-428.

CHAPTER 5: Physically sourced climatic changes and their implications on future flow projections

5.1 Introduction

Global climate change is expected to play a key role in shaping future climatic and hydrologic regimes. Global Circulation Models (GCMs) are the major tools that are majorly used to model these changes under the influence of future greenhouse gas emission scenarios. Raw GCM data has a typical spatial resolution of 1° x 1° which is inadequate for performing hydrological modeling at a catchment scale. To utilize GCM based climate projections for hydrological modelling, they need to be downscaled using either statistical or dynamic downscaling methods. A comprehensive review of different downscaling methods used in the past is provided in Schoof (2013).

Statistical downscaling methods estimate local climate by building a statistical relationship between large scale climatic or atmospheric variables with locally observed climatic data. On the other hand dynamic downscaling methods use large scale boundary conditions provided by a GCM as input into a Regional Climate Model (RCM) to simulate higher resolution data in a physically based way. They have been found to be useful for making future hydrologic predictions (Frost et al. 2011; Chiew et al. 2010; King et al. 2015; Srivastav and Simonovic 2014). For instance Chiew et al. (2010) used a scaling model, an analog based model, two regression based models and one dynamic downscaling method to downscale projections from three GCMs and found that statistical downscaling methods are able to simulate historical flow equally well as compared to the dynamical downscaling method. For this reason, many climate change impact assessment studies on flow and flooding regimes have

been performed by downscaling GCM projections from statistical downscaling methods (Gaur and Simonovic 2015; Schneider et al. 2012; Kingston and Taylor 2012; Grillakis et al. 2012; te Linde et al. 2010; Leander and Buishand 2007).

One common drawback identified in statistical downscaling methods is that they are not physically based. In Gaur and Simonovic (2016a) a Physical Scaling (SP) method is introduced which incorporates regional physical characteristics like elevation and land-cover into the statistical downscaling procedure. In Gaur and Simonovic (2016b) the method is extended to incorporate the physical characteristics of areas surrounding the reference location into the downscaling process as well. This method is referred to as SP method with Surrounding pixel information (SPS) method. Both methods have been tested extensively for their ability to downscale climate model based surface temperature, air temperature and precipitation data in Gaur and Simonovic (2016a; b; c).

Although hydrologic impact of future land-cover has been quantified in previous studies (Dwarakish and Ganasari 2015; Tejeda et al. 2014; Thanapakpawin et al. 2006), climatic influences of land-cover and associated impact on flows have not been quantified in the past. The inclusion of physical parameters into the downscaling process provides an opportunity to quantify future climatic changes as introduced by changes in physical characteristics of the location of interest within a statistical downscaling framework. The aim of this study is to quantify the magnitude and significance of these physically driven climatic changes. Further their impact on catchment outflow and flooding magnitudes is also analyzed.

Rest of the paper is organized as follows. A description of the study region is provided in section 5.2 which is followed by data used in section 5.3. A description of the models and

methods used in this study is provided in section 5.4, followed by a summary of the methodology we adopt in this study in section 5.5. This is followed by a discussion of the results in section 5.6 and conclusions in section 5.7.

5.2 Study region

Four catchments with HYDAT IDs: 05EG006, 05EG008, 05MC004 and 11AF005 are selected for analyses in this study. HYDAT is a flow database maintained by the Water Survey of Canada (WSC). More information about the data collected is provided in section 5.3. To calibrate the downscaling models used in this study, data are gathered from the entire southern Saskatchewan region. The physiographic setting of the study region as well as the location of selected catchments is shown in Figure 5.1. The black dots represent locations of precipitation and temperature gauging stations located within this region. The selected region is representative of Canadian prairies and is characterized by diverse topography and land-cover. Elevation varies between 240 masl to 1389 masl across the study region. Further all land-cover classes identified in the University of Maryland (UMD) classification scheme (summarized in Table 5.1) are found to be present within the study region. An analysis of MODerate-resolution Imaging Spectroradiometer (MODIS) land-cover data (MCD12Q1) between 2006 and 2013 suggests that land-cover classes: Cropland, Evergreen Needle-leaf Forest, Grassland and Mixed Forest constitute approximately 90% of the study region.
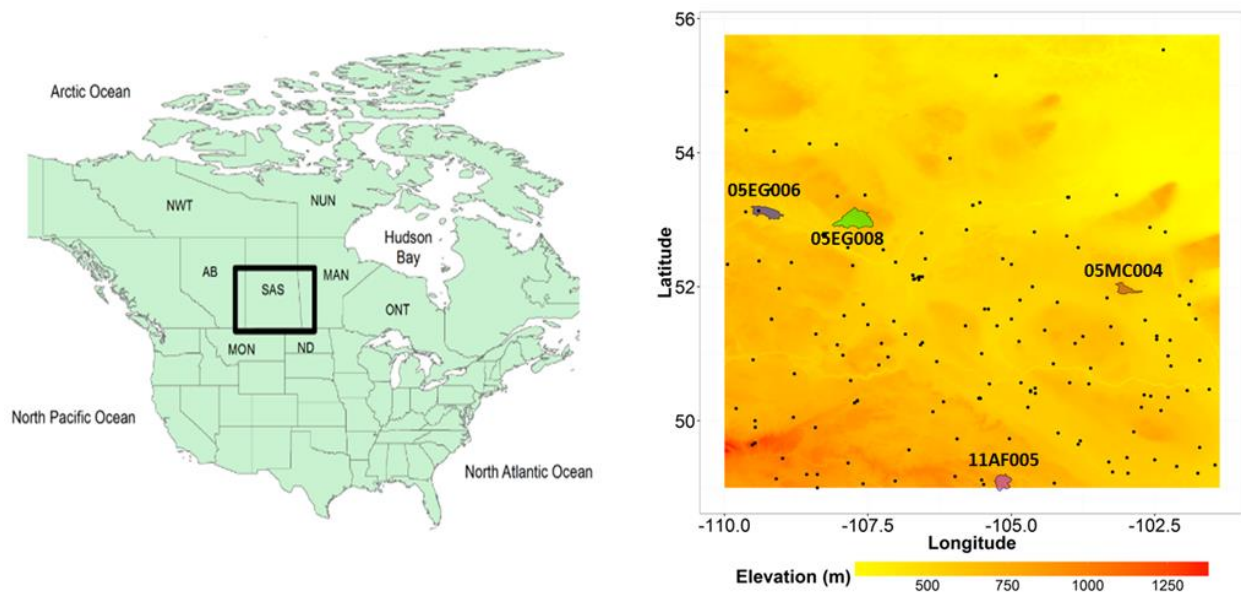
Figure 5.1. The geographic and political settings of the region of interest. Catchments selected are shown as blue, green, brown and purple shapes.

Table 5.1. Land-cover classes identified in the UMD classification system. Abbreviations used for different land-cover classes in this study are provided within brackets.

S.No	UMD classes
1	Water (W)
2	Evergreen Needleleaf Forest (ENF)
3	Evergreen Broadleaf Forest (EBF)
4	Deciduous Needleleaf Forest (DNF)
5	Deciduous Broadleaf Forest (DBF)
6	Mixed Forest (MF)
7	Closed Shrublands (CS)
8	Open Shrublands (OS)
9	Woody Savannas (WS)
10	Savannas (S)
11	Grasslands (G)
12	Croplands (C)
13	Urban and Built-up (UB)
14	Barren or Sparsely Vegetated (BSV)

The region is land-locked and encompasses many small lakes and rivers. It experiences a continental climate. Large fluctuations in temperature are observed owing to the land-locked location of the region. The region receives almost two-thirds of its precipitation during the summer season, which usually occurs due to large scale convective and cyclonic systems. Significant spatial variability in precipitation is also observed across this region. Snow cover plays a critical role in shaping the hydro-meteorology of the region as this region stays snow-covered almost six months a year (Encyclopedia of Saskatchewan 2016).

Monthly variations in the flows recorded at the selected discharge gauging stations are presented in Figure 5.2. It is evident from the figure that very high mean and extreme flows are observed in the months: April and May as compared to other months highlighting the role of snowmelt towards generating runoff in these catchments.
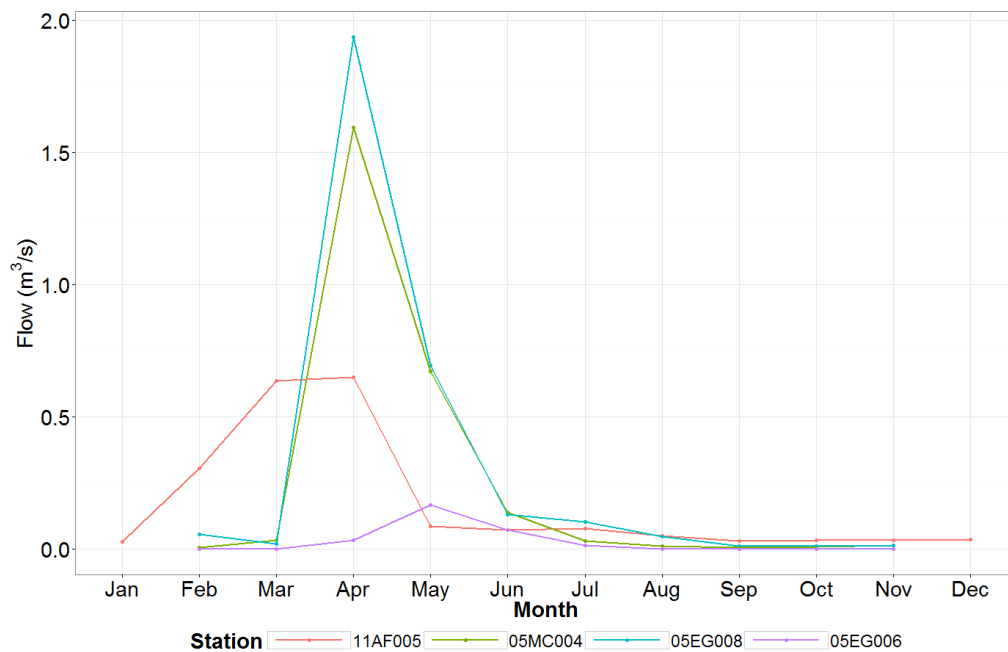
Figure 5.2. Monthly variations in flow mean across the four catchments considered for the analysis. Flow data for the period mentioned in Table 5.2 are used to derive the variations.

5.3 Data used

The following data has been used in this study:

SRTM elevation data: The National Aeronautics and Space Administration (NASA) Shuttle Radar Topographic Mission (SRTM) elevation product (Jarvis et al. 2008) is used in this study. This data has a spatial resolution of 90 m.

MODIS land-cover data: The MODIS land-cover data product (MCD12Q1) in UMD classification scheme is used in this study (LP DAAC 2001). A list of land-cover classes identified in UMD classification scheme is provided in Table 5.1. The data is available in 500 m spatial resolution at annual timesteps. Land-cover data for the period 2006-2012 is selected for analysis. Land-cover for the year 2013 is considered to be the same as that of year 2012 since data for that year is not available from MODIS data repository. This is a reasonable assumption since land-cover in the study region has not changed drastically at annual time-steps in the past (Gaur and Simonovic 2016a).

ANUSPLIN precipitation (ppt), maximum temperature (T_{max}) and minimum temperature (T_{min}) data: ANUSPLIN is a gridded climate data developed by applying thin plate spline smoothing algorithms on gauged climate records to obtain a continuous gridded dataset of 10 Km spatial resolution (Hopkinson et al. 2011; Hutchinson et al. 2009). In this study ANUSPLIN data for the period 1961-2013 located within the study region has been used.

Gauged daily precipitation and air temperature data: Daily air temperature and precipitation data gauged at 170 locations within the study region over the period 2006-2013 is acquired from Environment Canada. The data can be accessed at: http://climate.weather.gc.ca/. The distribution of gauging stations across the study region is shown in Figure 5.1. Using MODIS

land-cover data it is found that these gauging stations are associated with UMD land-cover classes: C, DNF, G, MF, OS, S, UB, WS, ENF, DNF and EBF.

HYDAT flow data: Daily flow data for the four selected catchments is collected from the HYDAT database maintained by the Water Survey of Canada (WSC). Flow data available between the period 1961 and 2013 is collected. A summary of the flow data length available at each flow gauging station, drainage area, mean elevation and HYDAT ID is provided in Table 5.2. All the selected catchments are non-regulated catchments so the observed changes in flow are driven by changes in physical characteristics of the catchment as well as external climatic forcings only.

Table 5.2. List of catchments selected for the analysis.

HYDAT ID	Name	Drainage area (Km²)	Elevation (masl)	Data length
05EG006	Birling creek near Paynton	426	593	1970-1992
05EG008	Page creek near Iffley	921	673	1971-1987
05MC004	Conjuring creek near Preeceville	255	594	1965-1992
11AF005	Beaver creek at international boundary	387	773	1977-1995

NARR climatic data: Daily precipitation rate and near surface air temperature data is acquired for the period 2006-2013 from NARR data repository (Mesinger et al. 2006). These data have an approximate spatial resolution of 32 Km and is prepared by forcing the National Center for Environmental Prediction (NCEP) Eta model with the Regional Data Assimilation System (RDAS).

Future land-cover projections: Future land-cover data for the period 2014-2100 for climate models listed in Table 5.3 and emission scenarios: RCP 2.6 and RCP 8.5 are taken from Gaur

and Simonovic (2016a). In Gaur and Simonovic (2016a) the 500 m future annual land-cover data is generated by downscaling and reclassifying the future harmonized land-use projections discussed in Hurtt et al. (2011). Future land-cover data is available in UMD classification system (Table 5.1).

Future GCM climatic data: GCM based daily air temperature and precipitation projections for the period 2006-2100 are collected from the Coupled Model Intercomparison Project-Phase 5 (CMIP5) data repository (Taylor et al. 2012). Data corresponding to climate models listed in Table 5.3 and for two Representative Concentration Pathways (RCPs): RCP 2.6 and RCP 8.5 are acquired. The choice of climate models is made based on the availability of future land-cover data as developed in Gaur and Simonovic (2016a).

Table 5.3. List of GCMs considered for the analysis (Flato et al. 2013).

GCM	Model	Resolution	Source
1	IAP-FGOALS	$1.66° \times 2.81°$	Institute of Atmospheric Physics, Chinese Academy of Sciences, China
2	MRI-CGCM3	$1.08° \times 2.16°$	Meteorological Research Institute, Japan
3	NorESM1-M	$2° \times 2°$	Norwegian Climate Centre, Norway

5.4 Models and methods used

Downscaling and hydrologic models used in this study as well as the statistical distribution used to fit peak flow data are presented in this section.

5.4.1 Downscaling models

A list of SP method based downscaling models is provided in Table 4. A brief description of these models is provided here however the reader is referred to Gaur and

Simonovic (2016b; c) for a more elaborate description of the downscaling models considered in this study.

SP method based models

In SP model, locally observed climate data is modelled using interpolated large scale climate model derived data as well as other explanatory variables representing land-cover, elevation properties of the location of interest. The predictor and predictant variables are linked using Generalized Additive Model (GAM) regression relationship. SP model approach to downscale air temperature can be mathematically expressed as:

$$g(T_{obs}) = B_0 + f_1(T_{mod}) + f_2(E_p) + f_3(LC_p) \qquad (5.1)$$

Where, *P* and *T* denotes precipitation and air temperature respectively, *E* denotes elevation (masl) of the reference pixel, *LC* denotes categorical land-cover variable associated with the reference pixel and *B* denote regression parameters. Subscript *obs* and *mod* describe if the climatic data is observed or model based respectively. In the case of climate models, P_{mod} denotes bilinearly interpolated climate model data at the reference pixel. Subscript *p* indicates that the data used is a pixel scale data whereas subscript *wet* denotes values on wet days only (i.e. days with more than 0.1 mm of precipitation). Variables *g* and *f* represent the link function and smoothing functions respectively. In this study, the smooth function is fit using penalized likelihood maximization algorithm. The penalized likelihood maximization algorithm is a variant of maximum likelihood estimation algorithm and applies a tradeoff between model fit wiggliness and goodness of fit by incorporating a penalty function (Wood 2000).

In case of precipitation the model involves a two-step process of predicting precipitation occurrence using a logistic regression model and a wet day precipitation amounts model using GAM regression model. SP model approach to downscale precipitation can be mathematically expressed as:

$$\ln\left(\frac{P_{obs}}{1-P_{obs}}\right) = B_0 + B_1 P_{mod} + B_2 E_p + B_3 LC_p \qquad (5.2)$$

$$g(P_{obs,wet}) = B_0 + f_1(P_{mod,wet}) + f_2(E_p) + f_3(LC_p) \qquad (5.3)$$

Where, notations have previously defined meanings. Additionally subscript *wet* denotes values on wet days only (i.e. days with more than 0.1 mm of precipitation). One variant of the SP model is also considered in this study. This model ignores both land-cover and elevation as predictors from SP model formulation and considers only interpolated climate model data as predictor. This model is referred to as SP_LC_elev in this paper.

Table 5.4. List of SP method based models used for downscaling climate model projections in this study (After Gaur and Simonovic 2016c). Here CV denotes climate variable of interest, LC denotes land-cover of the reference pixel, E denotes elevation, NLC denotes land-cover of neighborhood pixels.

S.No	Model name	Predictors
1	SP	CV, LC, E
2	SP_LC_elev	CV
3	SPS3x3	CV, LC, E, NLC_{3x3}, NE_{3x3}
4	SPS5x5	CV, LC, E, NLC_{5x5}, NE_{5x5}
5	SPS7x7	CV, LC, E, NLC_{7x7}, NE_{7x7}
6	SPS9x9	CV, LC, E, NLC_{9x9}, NE_{9x9}

SP method with Surrounding pixel information (SPS) method based models

SPS method incorporates land-cover and elevation properties of the neighborhood pixels into the SP model formulation. The SPS model for downscaling air temperature can be mathematically expressed as:

$$g(T_{obs}) = B_0 + f_1(T_{mod}) + f_2(E_p) + f_3(LC_p) + f_4(Fr_{W,s}) + ... + f_{21}(Fr_{BSV,s}) + f_{22}(R_{E,s}) \quad (5.4)$$

Where, symbols have similar meanings as explained above. Predictors $Fr_{W,s}$, ...$Fr_{BSV,s}$ represent the fraction of total area surrounding the reference pixel that is occupied by *Water,....Barren and Sparsely Vegetated* land-cover classes respectively. The value of predictors: $Fr_{W,s}$, ...$Fr_{BSV,s}$ is between 0 and 1 and they add up across all neighborhood land-cover classes to give a value of 1. Neighborhood elevation information is incorporated by including a predictor $R_{E,s}$ which represents the ratio between reference pixel elevation and mean elevation of pixels surrounding the reference pixel.

For precipitation, again the SPS downscaling method involves two steps of forming a precipitation occurrence and wet day precipitation amounts model. The two steps involved in SPS method can be mathematically expressed as:

$$\ln\left(\frac{P_{obs}}{1 - P_{obs}}\right) = B_0 + B_1 P_{mod} + B_2 E_p + B_3 LC_p + B_4 Fr_{W,s} + ... + B_{21} Fr_{BSV,s} + B_{22} R_{E,s} \quad (5.5)$$

$$g(P_{obs,wet}) = B_0 + f_1(P_{mod,wet}) + f_2(E_p) + f_3(LC_p) + f_4(Fr_{W,s}) + ... + f_{21}(Fr_{BSV,s}) + f_{22}(R_{E,s}) \quad (5.6)$$

Where, symbols have similar meanings as explained above.

SPS method based models are formulated at four neighborhood scales: 3x3, 5x5, 7x7 and 9x9 (represented as *s* in equations 5.4, 5.5 and 5.6) in this study. The neighborhood scales considered in this study are shown in Figure 5.3. In the figure, reference pixel is shown in red while the 3x3, 5x5, 7x7 and 9x9 neighborhood scales are shown in orange, yellow, green and blue respectively. Larger neighborhood scales are considered to be inclusive of smaller spatial scales which for instance imply that neighborhood scale 5x5 will encompass the pixels corresponding to neighborhood scale: 3x3 and additional yellow pixels. These neighborhood scales have been chosen in previous studies (Gaur and Simonovic 2016b; c; Verburg et al. 2004) and have been adopted in this study as well.
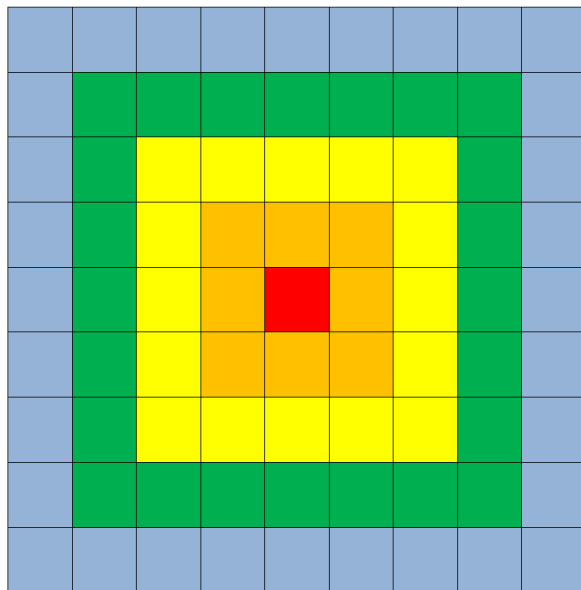
Figure 5.3. Neighborhood scales considered in this study. The 3x3, 5x5, 7x7 and 9x9 neighborhood scale is shown in orange, yellow, green and blue respectively while the reference pixel is shown in red (after Gaur and Simonovic 2016b).

5.4.2 Sacramento Soil Moisture Accounting model: In the Sacramento model (Burnash 1995) flow is generated by distributing precipitation falling at a location to overland flow, interflow and baseflow components accounting for losses due to evapotranspiration and interception. Groundwater movement is modelled by considering upper zone and lower zone storages. Runoff is contributed by five different processes: (1) direct runoff from permanent and temporary impervious areas, (2) surface runoff due to precipitation occurring at a rate faster than percolation and interflow that take place when both upper zone storages are full, (3) interflow resulting from the lateral drainage of a temporary free water storage, (4) supplemental base flow, and (5) primary base flow. The model has 13 free parameters which are optimized.

Routing in this model is performed using exponential form of unit hydrograph with explicit slow and quick flow components. The routing scheme involves 3 free parameters which are optimized. A list of hydrologic model and routing parameters are presented in Table D1. Snow-melt is modelled offline using a temperature index modelling approach discussed in Walter et al. (2005). Model is calibrated to optimize an objective function which is a weighted sum of coefficient of determination (R^2) and relative bias. The shuffled complex evolution global optimization method (Duan et al. 1993) followed by a local optimization method (Nelder and Mead 1965) with multi-start options is used to calibrate the model. A lumped version of this hydrologic model is considered sufficient for this study because the catchments considered are small (drainage area < 1000 Km^2) and can be modelled using a lumped model depicting essential hydrological processes within its framework.

5.4.3 Generalized Extreme Value (GEV) distribution

In this study flow extremes are analyzed by using a block maxima approach. This approach is considered suitable for our analysis because the data being fitted is of sufficient length (86 years) needed to derive reliable distributional parameters for the function used to fit extreme values. In the block maxima method, the independent and identically distributed (iid) samples are chosen from a block of data falling within a selected duration of time (typically taken as a year) and are fitted using Generalized Extreme Value (GEV) distribution. The mathematical form of GEV distribution is expressed in equation 5.7.

$$G(z) = \exp\left[-\left\{1 + \xi\left(\frac{z - \mu}{\sigma}\right)\right\}^{-1/\xi}\right] \tag{5.7}$$

Where, *G*, *E* and *P* denote intensity measures for the GEV distribution. Parameters μ, σ and ξ denote location, scale and shape parameters respectively. In this study parameters μ, σ and ξ are estimated using L-moments method (Hosking 1990).

5.5 Analyses performed

Following analyses are conducted in this study:

5.5.1 Calibration of downscaling models: Downscaling models listed in Table 5.4 are calibrated over the period 2006-2013 using gauged climate data as predictant variable and NARR based gridded climate data and other physical variables as predictor variables. Separate sets of parameters are derived for snow-covered and snow-free

months. In this study period from April to September is considered as snow-free months and October to March is considered as snow-covered months.

5.5.2 Downscaling of future climate model projections: The calibrated downscaling models are thereafter used to downscale future GCM maximum temperature, minimum temperature and precipitation projections for the period 2014-2100 at all MODIS grid cells falling within each catchment.

5.5.3 Calibration of hydrologic model: The Sacramento hydrologic model is calibrated for each catchment using discharge data available between 1961 and 2013 at each discharge gauging station. The available daily discharge data length at each catchment is provided in Table 5.2. Since very few temperature and precipitation gauging stations are found to be located within each selected catchment, ANUSPLIN based precipitation, maximum temperature, and minimum temperature gridded data are used as climatic forcing in the hydrologic model while calibrating them.

5.5.4 Prediction of future discharge from catchments: Calibrated Sacramento model is thereafter used to predict future flows from the catchments. Downscaled maximum temperature, minimum temperature and precipitation projections obtained from different downscaling models are used as inputs into the hydrologic model to derive future catchment discharges.

5.5.5 Flood frequency analysis of projected flows: Extreme value analysis is performed on future projected flow data. Five different return periods: 2, 5, 10, 25 and 100 years are chosen to analyze the impacts on flooding events of different intensities.

5.5.6 Fingerprinting difference in projections to changes in physical characteristics: Future projected precipitation, mean temperature (average of maximum and minimum temperature), mean and extreme discharge at each catchment are analyzed to identify the impact of land-cover driven climatic changes on these variables. For doing so projections made by downscaling model: SP_LC_elev (which only considers GCM data as predictor) are compared with those obtained from models: SP, SPS3x3, SPS5x5, SPS7x7, SPS9x9 (which consider other physical based parameters together with GCM data as predictors).

5.6 Results and discussion

Future temperature and precipitation projections made by different downscaling methods considered in this study are presented in Figures 5.4 and 5.5 for all four catchments. Average precipitation across the catchment is projected to increase in all four catchments. It is projected that future precipitation will increase on average at the rate of 0.001 mm/year ($p = 0.001$), 0.001 mm/year ($p < 0.001$), 0.001 mm/year ($p = 0.003$) and 0.002 mm/year ($p < 0.001$) for catchments 05EG006, 05EG008, 05MC004 and 11AF005 respectively. In the case of mean temperature, it is projected to increase at an average rate of 0.03 K/year ($p < 0.001$) across all four catchments.
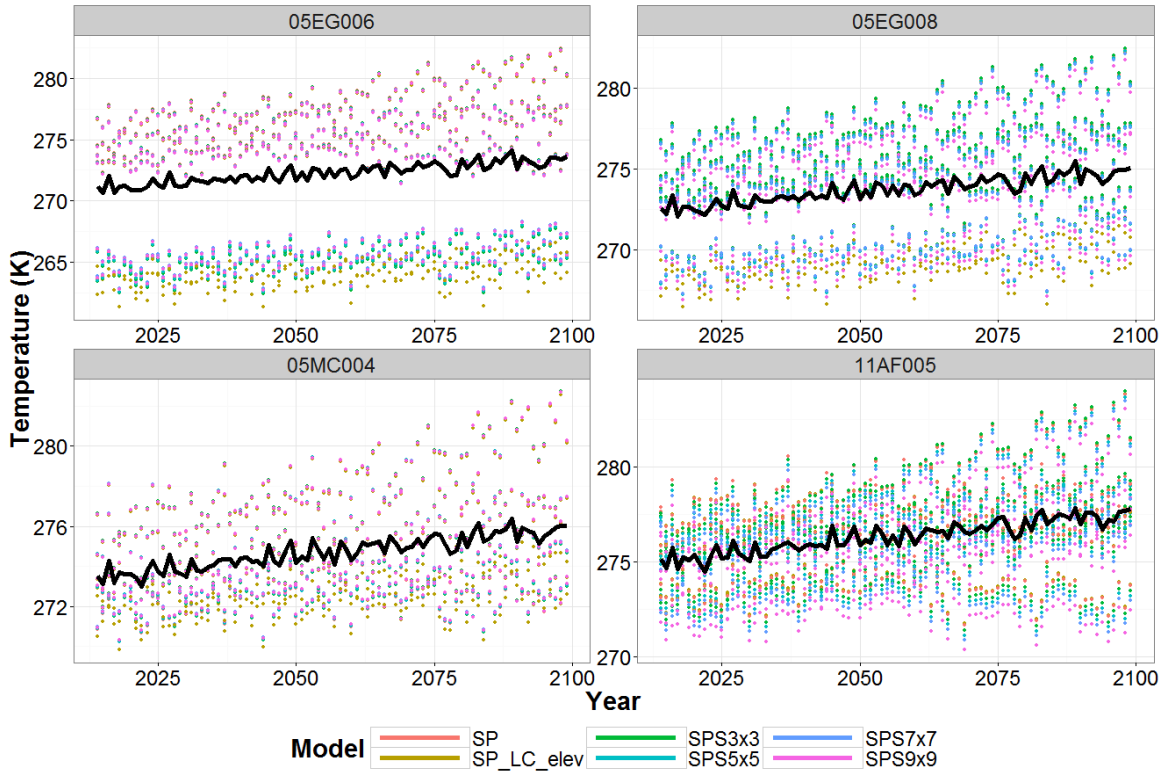
Figure 5.4. Catchment averaged temperature for the selected catchments over the period 2014-2100 from different downscaling models. Average trendline from all downscaling models is shown in bold black.
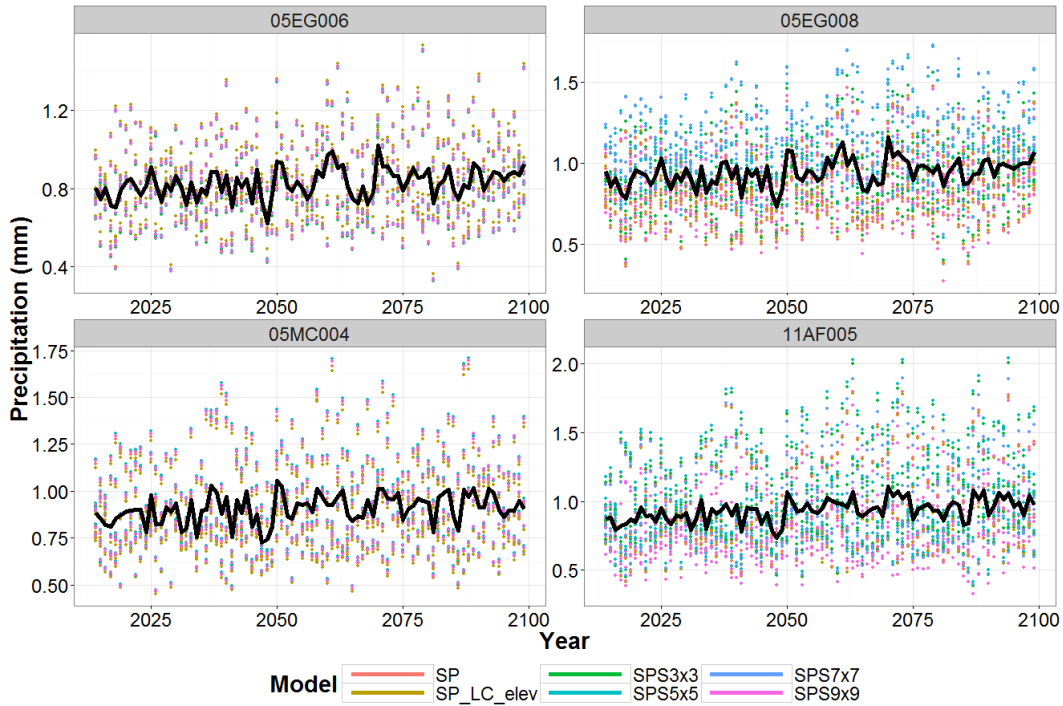
Figure 5.5. Catchment averaged precipitation for the selected catchments over the period 2014-2100 from different downscaling models. Average trendline from all downscaling models is shown in bold black.

Sacramento hydrologic model is calibrated for the four catchments. Calibration results for the model in each catchment are provided in Table 5.5. The Nash-Sutcliffe efficiency (Nash and Sutcliffe 1970) and correlation between observed and modelled streamflow is analyzed. Since we are only interested in looking at the difference in projections from a set of climatic projections, the model calibration is considered to be satisfactory for our study. The entire flow data length is considered while calibrating the model to obtain a set of robust hydrologic parameters which can be used to predict future flows. The downscaled precipitation and temperature from each downscaling model is used as input into the hydrologic model to predict future flows from each catchment. Unlike precipitation and temperature, future

projected flows don't show statistically significant future trends in any of the catchments. This is also evident from Figure 5.6 where yearly averaged future flow projections from each downscaling model are shown.

Table 5.5. Calibration results of Sacramento model for each catchment.

HYDAT ID	Name	NSE	Correlation
05EG006	Birling creek near paynton	0.65	0.83
05EG008	Page creek near iffley	0.58	0.79
05MC004	Conjuring creek near preeceville	0.58	0.78
11AF005	Beaver creek at international boundary	0.50	0.70
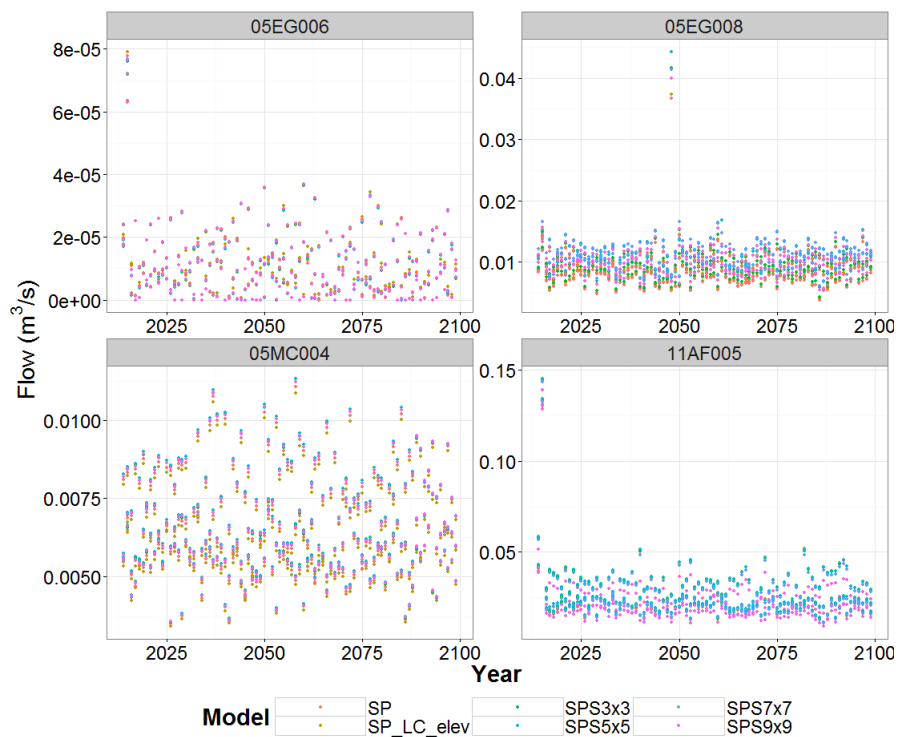
Figure 5.6. Flows generated from each catchment over the period 2014-2100 from different downscaling models.

To explore if the consideration of catchment physical characteristics effect climate and hydrologic variables, downscaled projections obtained from SP_LC_elev downscaling model are compared with the projections obtained from all other downscaling models using Wilcoxon signed rank test (Wilcoxon 1945). Wilcoxin signed rank test can compare two paired samples of data without assuming normality in their distributions. Tests are conducted on catchment averaged projected precipitation, temperature and flow data obtained from all GCM-RCP combinations considered in this study. A summary of the results is provided in Figures 5.7-5.9. The *p* values presented here for different downscaling model are obtained by considering projections derived from SP_LC_elev downscaling model as reference data and each of the other models as test data. They correspond to the rejection of null hypothesis that both samples are taken from the same population. The highlighted cells indicate cases where a statistically significant rejection of the null hypothesis is obtained. It can be seen that for all three variables, projections based on SP_LC_elev model is statistically different from that obtained from other models at 0.05 significance level. In the case of precipitation, models: SPS3x3 and SPS5x5 are found to force highest number of statistically significant changes (in 100% of the total scenarios considered), followed by SPS7x7 (92%), followed by SPS9x9 (88%) and followed by SP (50%) model. In the case of temperature, model: SPS9x9 is found to bring most number of statistically significant changes (in 54% of the total scenarios considered), followed by SPS7x7 (38%), followed by models: SP, SPS3x3 and SPS5x5 (33%). In the case of flow projections, model: SPS5x5 and SPS7x7 are found to bring most number of statistically significant changes (in 92% of the total scenarios considered), followed by SPS3x3 (88%), followed by SPS9x9 (83%) and followed by SP (58%).
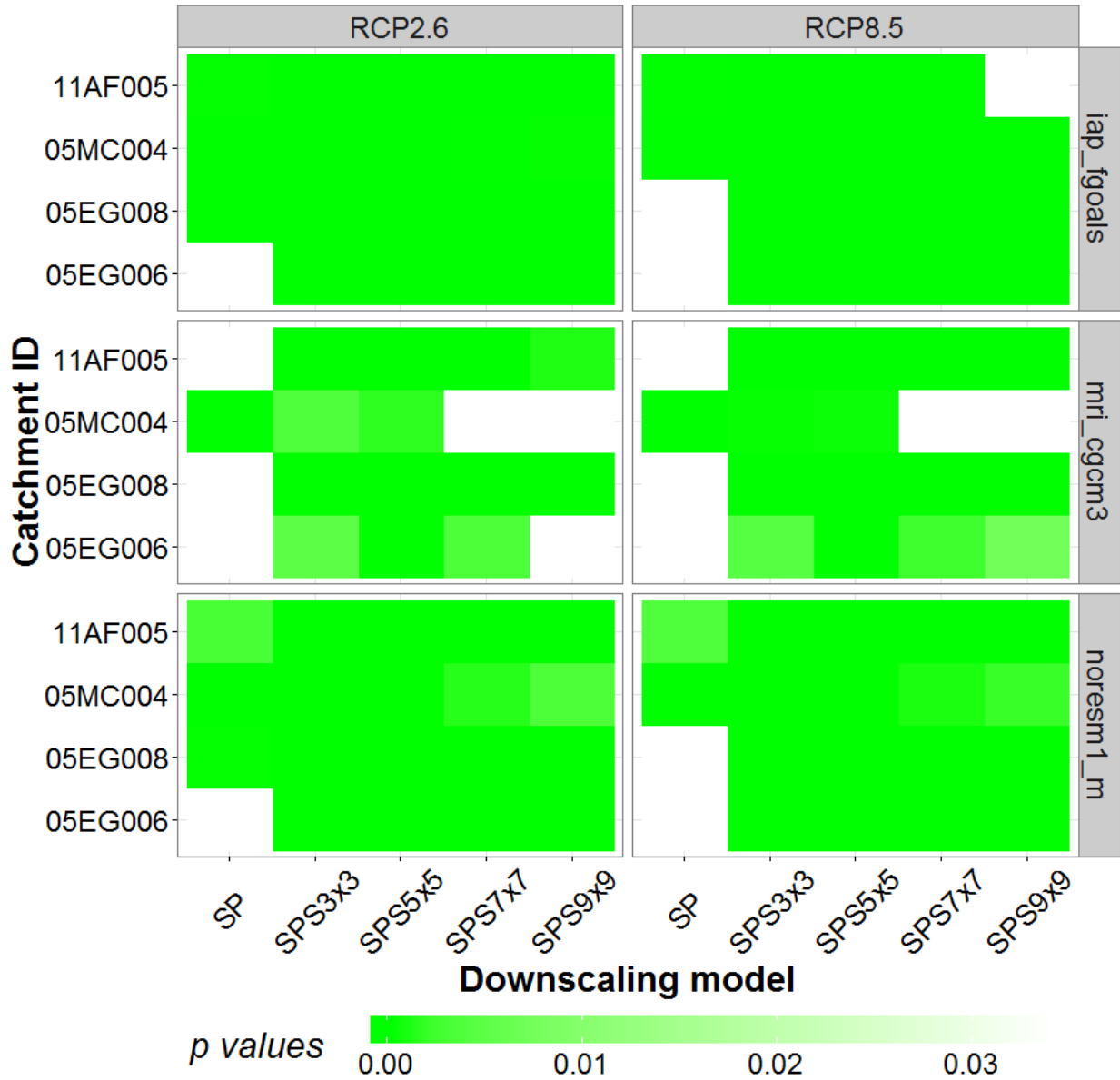
Figure 5.7. Probability of rejection of the null hypothesis that precipitation projections obtained from the downscaling model: SP_LC_elev and other downscaling models are from the same population.
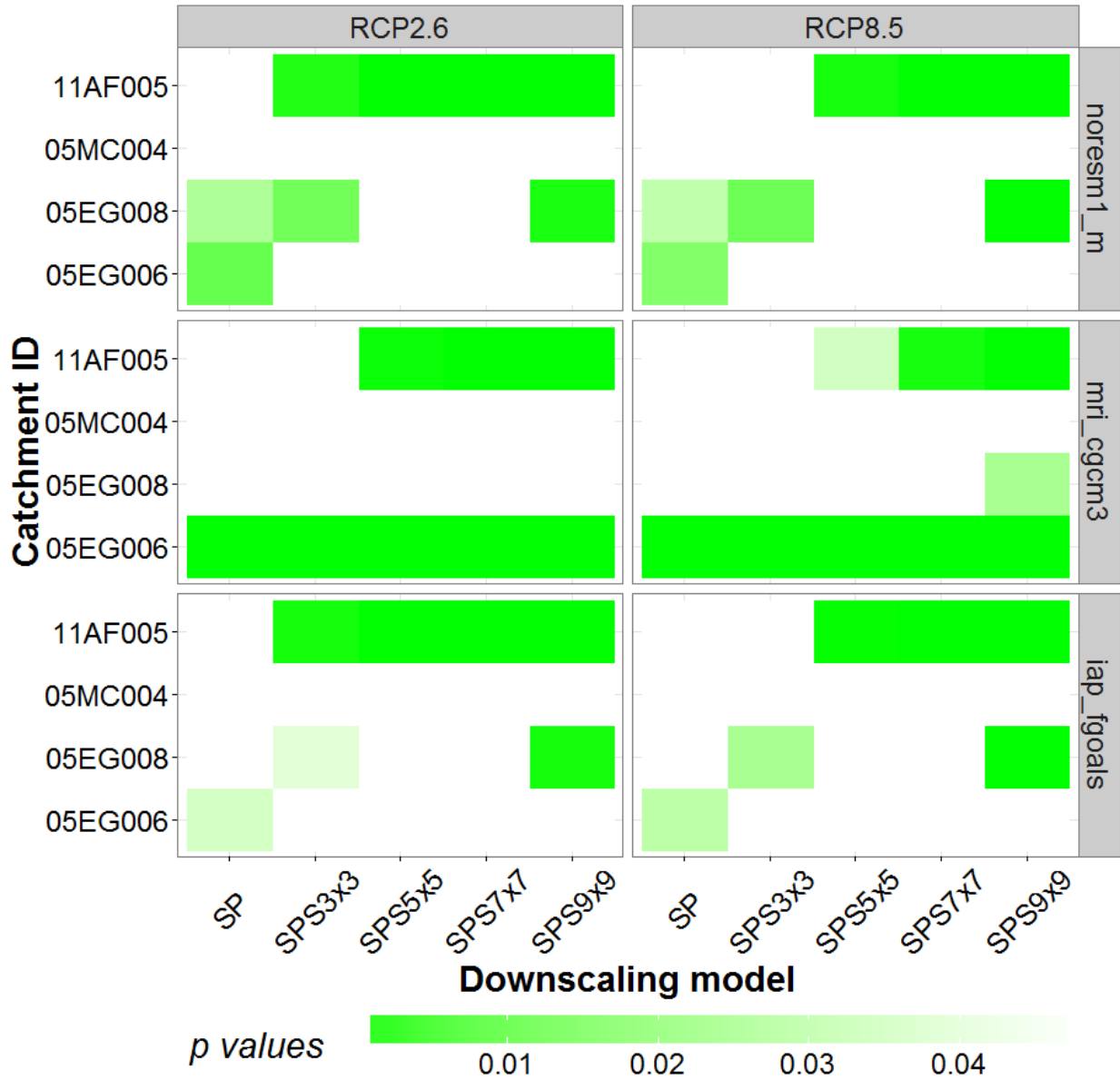
Figure 5.8. Probability of rejection of the null hypothesis that temperature projections obtained from the downscaling model: SP_LC_elev and other downscaling models are from the same population.
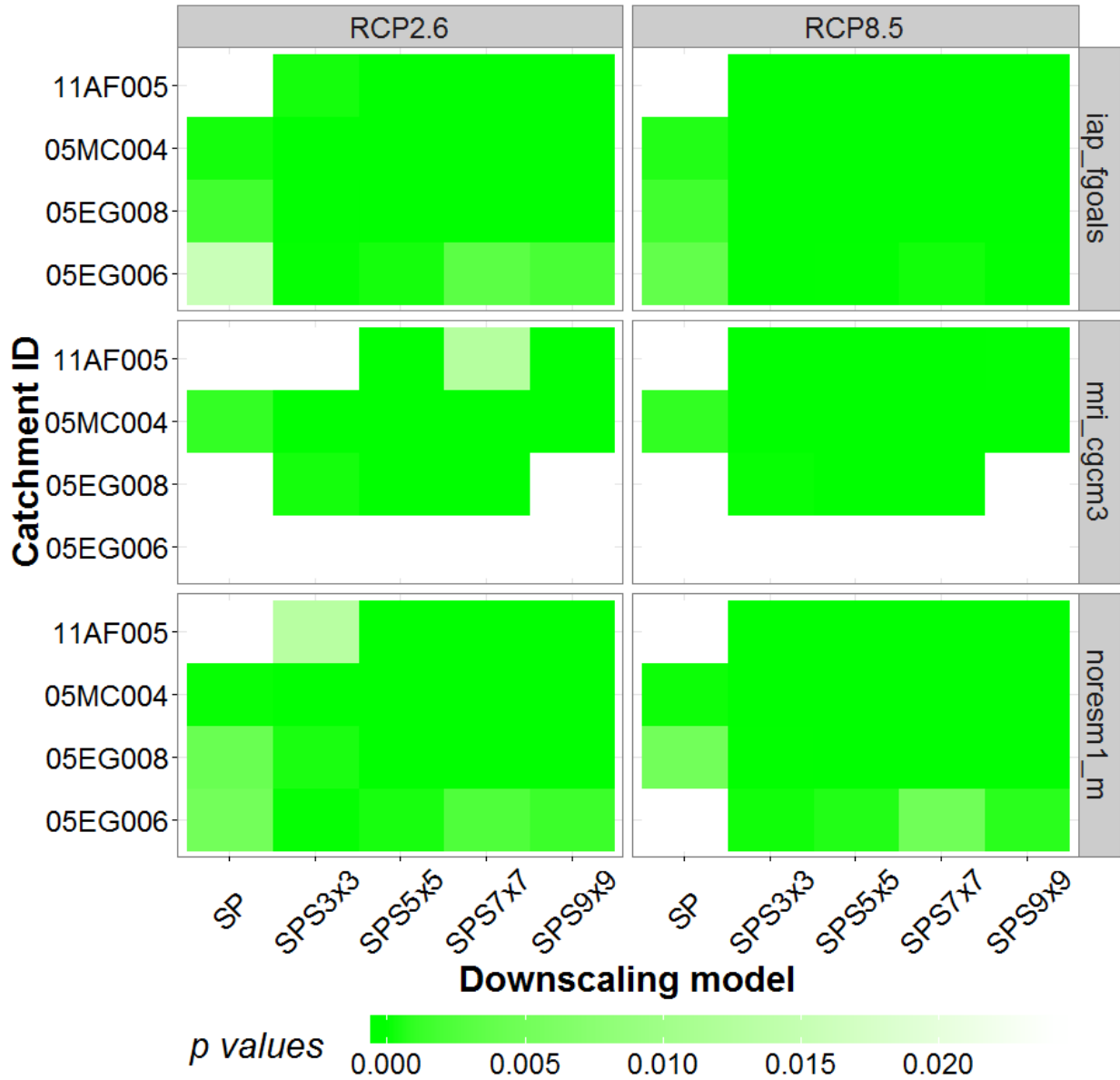
Figure 5.9. Probability of rejection of the null hypothesis that flow projections obtained from the downscaling model: SP_LC_elev and other downscaling models are from the same population.

The magnitude of Physically sourced changes (PSC) between SP_LC_elev and other downscaling models are summarized in Tables 5.6-5.8 and Figure 5.10 for each catchment considered in this study. Absolute increases in temperature and percent increases in precipitation and flow are used as an index to represent PSC magnitude. Further cells showing positive and negative changes are highlighted in red and blue respectively in Tables 5.6-5.8. It is found that among the four catchments considered, largest catchment averaged changes in precipitation correspond to the catchment: 05EG008 (12%), followed by 05MC004 (4%), followed by 11AF005 (3%) while an overall decrease of 4% is noted for catchment 05EG006. In the case of temperature, largest increases are projected for the catchment 05EG006 (0.5 K), followed by 05EG008 (0.3 K), followed by 05MC004 (0.2 K) while a decrease of 0.5 K is noted for the catchment 11AF005. In the case of flow, largest increases are observed for catchment 05EG008 (2%), followed by catchments 05MC004 and 11AF005 (1%) and followed by 05EG006 which shows almost no mean change.

Table 5.6. A summary of mean physically sourced changes for precipitation. Cells showing positive and negative changes are highlighted in red and blue respectively.

Catchment	GCM	RCP	SP	SPS3x3	SPS5x5	SPS7x7	SPS9x9
05EG006	iap_fgoals	RCP2.6	-3.1	-5.5	-4.5	-4.1	-4.9
		RCP8.5	-3.2	-5.8	-4.8	-4.3	-5.0
	mri_cgcm3	RCP2.6	-1.1	-2.4	-2.1	-1.7	-2.1
		RCP8.5	-0.9	-2.2	-1.9	-1.5	-1.9
	noresm1_m	RCP2.6	-3.3	-5.8	-4.8	-4.3	-5.1
		RCP8.5	-3.0	-5.6	-4.7	-4.2	-4.8
05EG008	iap_fgoals	RCP2.6	-2.1	6.1	29.8	32.5	9.2
		RCP8.5	-2.0	6.2	33.9	40.6	4.4
	mri_cgcm3	RCP2.6	-0.7	5.3	17.6	16.9	0.5
		RCP8.5	-0.6	5.4	19.3	20.3	-6.2
	noresm1_m	RCP2.6	-2.0	6.0	25.6	26.7	5.1
		RCP8.5	-2.0	3.6	30.4	34.5	-3.2
05MC004	iap_fgoals	RCP2.6	2.2	3.4	4.5	3.5	3.7
		RCP8.5	2.1	3.6	4.7	3.6	3.6
	mri_cgcm3	RCP2.6	1.8	3.1	3.9	3.1	3.2
		RCP8.5	1.8	3.2	4.1	3.1	3.2
	noresm1_m	RCP2.6	2.2	3.6	4.7	3.7	3.8
		RCP8.5	2.1	3.7	4.9	3.7	3.8
11AF005	iap_fgoals	RCP2.6	0.8	0.9	4.9	-0.3	-14.7
		RCP8.5	1.2	17.6	25.7	12.8	0.1
	mri_cgcm3	RCP2.6	0.4	0.5	2.3	-2.3	-16.2
		RCP8.5	0.8	10.9	13.8	5.6	-6.7
	noresm1_m	RCP2.6	0.5	0.4	3.8	-1.8	-15.2
		RCP8.5	1.1	15.0	21.6	9.8	-2.5

Table 5.7. A summary of mean physically sourced changes for temperature. Cells showing positive and negative changes are highlighted in red and blue respectively.

Catchment	GCM	RCP	SP	SPS3x3	SPS5x5	SPS7x7	SPS9x9
05EG006	iap_fgoals	RCP2.6	0.2	0.1	0.1	0.1	-0
		RCP8.5	0.2	0.1	0.1	0.1	+0
	mri_cgcm3	RCP2.6	1.3	1.1	1.2	1.5	1.4
		RCP8.5	1.3	1.1	1.3	1.6	1.5
	noresm1_m	RCP2.6	0.2	0.1	0.1	0.1	+0
		RCP8.5	0.2	0.1	0.1	0.1	0.1
05EG008	iap_fgoals	RCP2.6	0.1	0.2	+0	-0.1	-0.4
		RCP8.5	0.2	0.2	+0	-0.1	-0.5
	mri_cgcm3	RCP2.6	0.8	0.9	0.8	0.9	0.5
		RCP8.5	0.9	0.9	0.8	0.8	0.4
	noresm1_m	RCP2.6	0.2	0.2	+0	-0.1	-0.4
		RCP8.5	0.2	0.2	0.1	-0.1	-0.5
05MC004	iap_fgoals	RCP2.6	+0	0.1	+0	+0	+0
		RCP8.5	+0	0.1	0.1	+0	+0
	mri_cgcm3	RCP2.6	0.4	0.4	0.4	0.5	0.5
		RCP8.5	0.4	0.4	0.5	0.5	0.5
	noresm1_m	RCP2.6	0.1	0.1	0.1	0.1	0.1
		RCP8.5	0.1	0.1	0.1	0.1	0.1
11AF005	iap_fgoals	RCP2.6	-0.1	-0.4	-0.6	-0.9	-1.4
		RCP8.5	-0.1	-0.1	-0.4	-0.6	-1.1
	mri_cgcm3	RCP2.6	+0	-0.3	-0.5	-0.8	-1.3
		RCP8.5	+0	+0	-0.3	-0.5	-1.0
	noresm1_m	RCP2.6	-0	-0.3	-0.6	-0.8	-1.3
		RCP8.5	-0.1	-0	-0.4	-0.6	-1.0

Table 5.8. A summary of mean physically sourced changes for streamflow at catchment outlet.

Cells showing positive and negative changes are highlighted in red and blue respectively.

| Catchment	GCM	RCP	SP	SPS3x3	SPS5x5	SPS7x7	SPS9x9
05EG006	iap_fgoals	RCP2.6	-0	-0	-0	-0	-0
		RCP8.5	-0	-0	-0	-0	-0
	mri_cgcm3	RCP2.6	-0	-0	-0	-0	-0
		RCP8.5	-0	-0	-0	-0	-0
	noresm1_m	RCP2.6	-0	-0	-0	-0	-0
		RCP8.5	-0	-0	-0	-0	-0
05EG008	iap_fgoals	RCP2.6	-0.2	0.5	2.4	2.6	1.5
		RCP8.5	-0.2	0.5	2.8	3.4	2.0
	mri_cgcm3	RCP2.6	-0.1	0.5	1.7	1.6	0.8
		RCP8.5	-0.1	0.6	2.0	2.1	1.0
	noresm1_m	RCP2.6	-0.2	0.4	2.0	2.1	1.2
		RCP8.5	-0.2	0.5	2.3	2.6	1.4
05MC004	iap_fgoals	RCP2.6	0.3	0.5	0.6	0.5	0.5
		RCP8.5	0.3	0.5	0.7	0.5	0.5
	mri_cgcm3	RCP2.6	0.3	0.6	0.7	0.5	0.6
		RCP8.5	0.3	0.6	0.8	0.6	0.6
	noresm1_m	RCP2.6	0.3	0.5	0.6	0.5	0.5
		RCP8.5	0.3	0.5	0.6	0.5	0.5
11AF005	iap_fgoals	RCP2.6	0.2	0.4	1.0	-0.1	-3.0
		RCP8.5	0.3	4.1	5.5	2.7	0.1
	mri_cgcm3	RCP2.6	0.1	0.4	0.7	-0.7	-5.2
		RCP8.5	0.3	3.7	4.3	1.7	-2.0
	noresm1_m	RCP2.6	0.1	0.4	0.8	-0.4	-3.1
		RCP8.5	0.2	3.4	4.4	2.0	-0.4
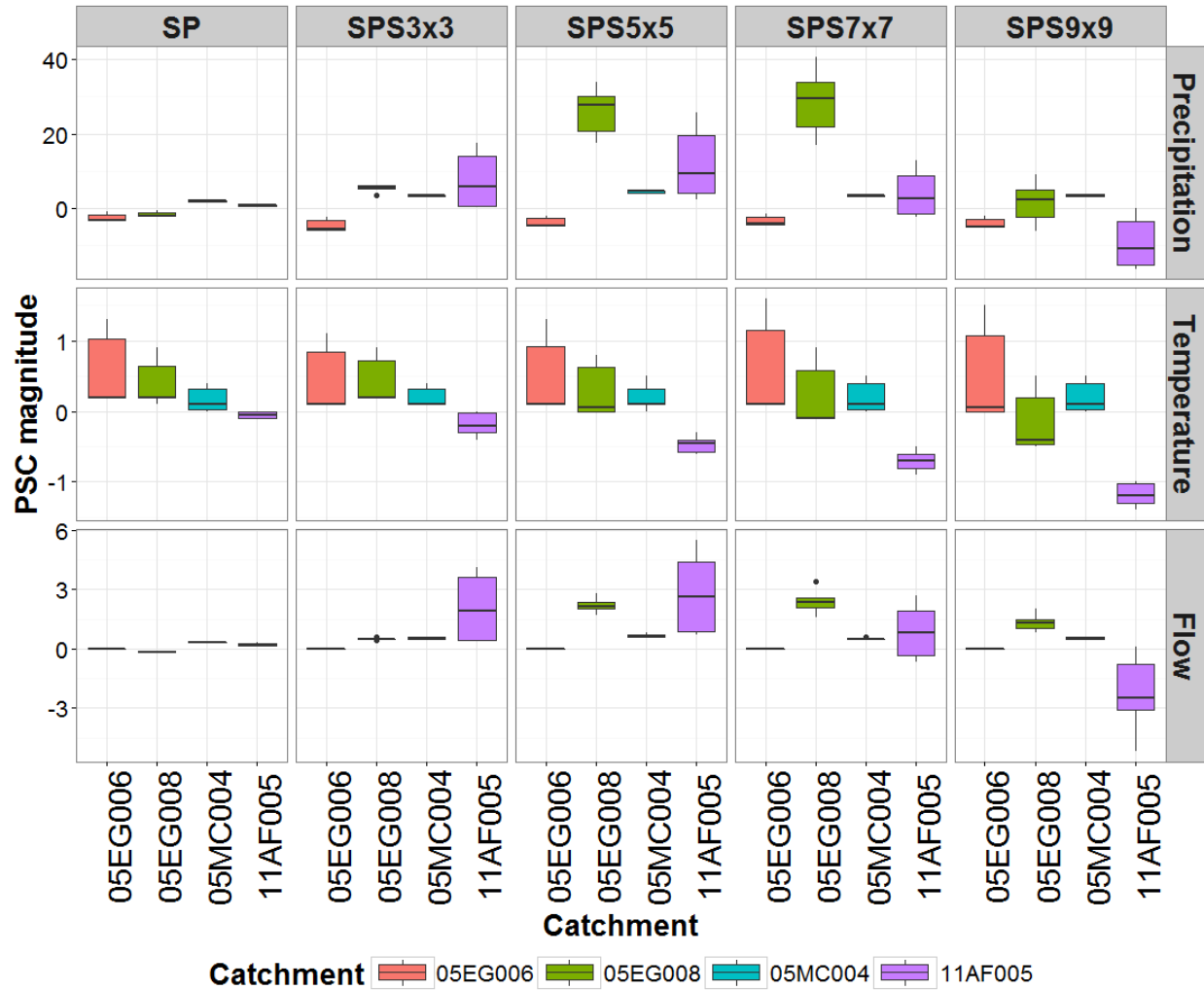
Figure 5.10. Differences in mean precipitation, temperature and flow projections from different downscaling models with respect to the model: SP_LC_elev. Here PSC magnitudes are provided in percentage for precipitation and flows while absolute differences are provided for temperature.

It must be noted that above mentioned PSC magnitudes are spatially averaged across the catchment area and temporally averaged over the period 2014-2100. Significant spatial and temporal variability in PSC magnitudes exists as shown in Figures 5.11 and 5.12. This is explored by analyzing the annual coefficient of variation (COV) values at each pixel located within the catchments. The COV is the percent ratio of sample standard deviation with sample mean.

The annual PSC magnitudes for precipitation and temperature are presented in Figure 5.11 for all catchments considered in this study. Boxplots of annual variations in the COV value are presented. For both variables, significant variations in PSC magnitudes are noted across the study region as well as over the 21st century. Larger variations are obtained for precipitation as compared to temperature. Highest variations are obtained for the catchment 11AF005, followed by 05MC004, followed by 05EG006, and followed by 05EG008. The COV for flow values for the period 2014-2100 are shown in Figure 5.12. It can be noticed by the annual COV values that PSC values again vary significantly over time and even within a year owing to projections from different GCMs, RCPs and downscaling models. Overall largest variations in PSC flow values are obtained for the catchment 05EG006 (which showed almost negligible mean PSC values), followed by catchment 05MC004, followed by catchment 05EG008 and followed by catchment 11AF005.
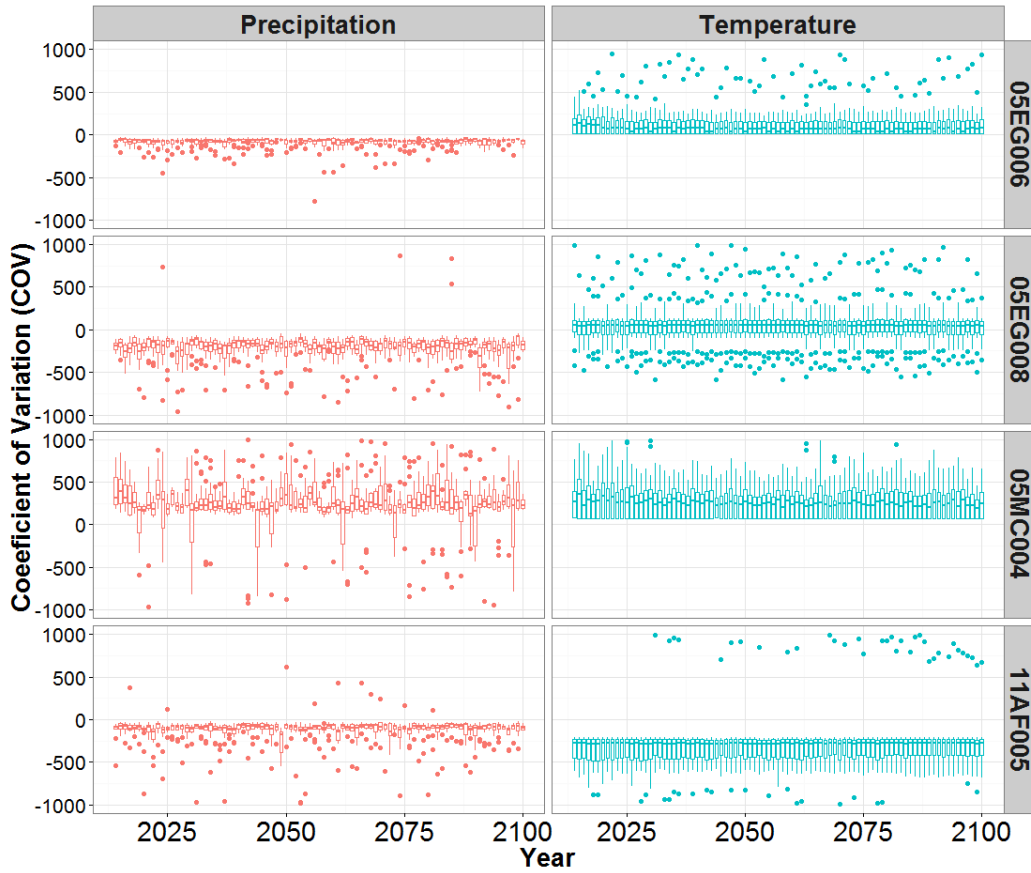
Figure 5.11. Annual boxplots showing spatial and temporal variation in PSC values for precipitation and temperature.
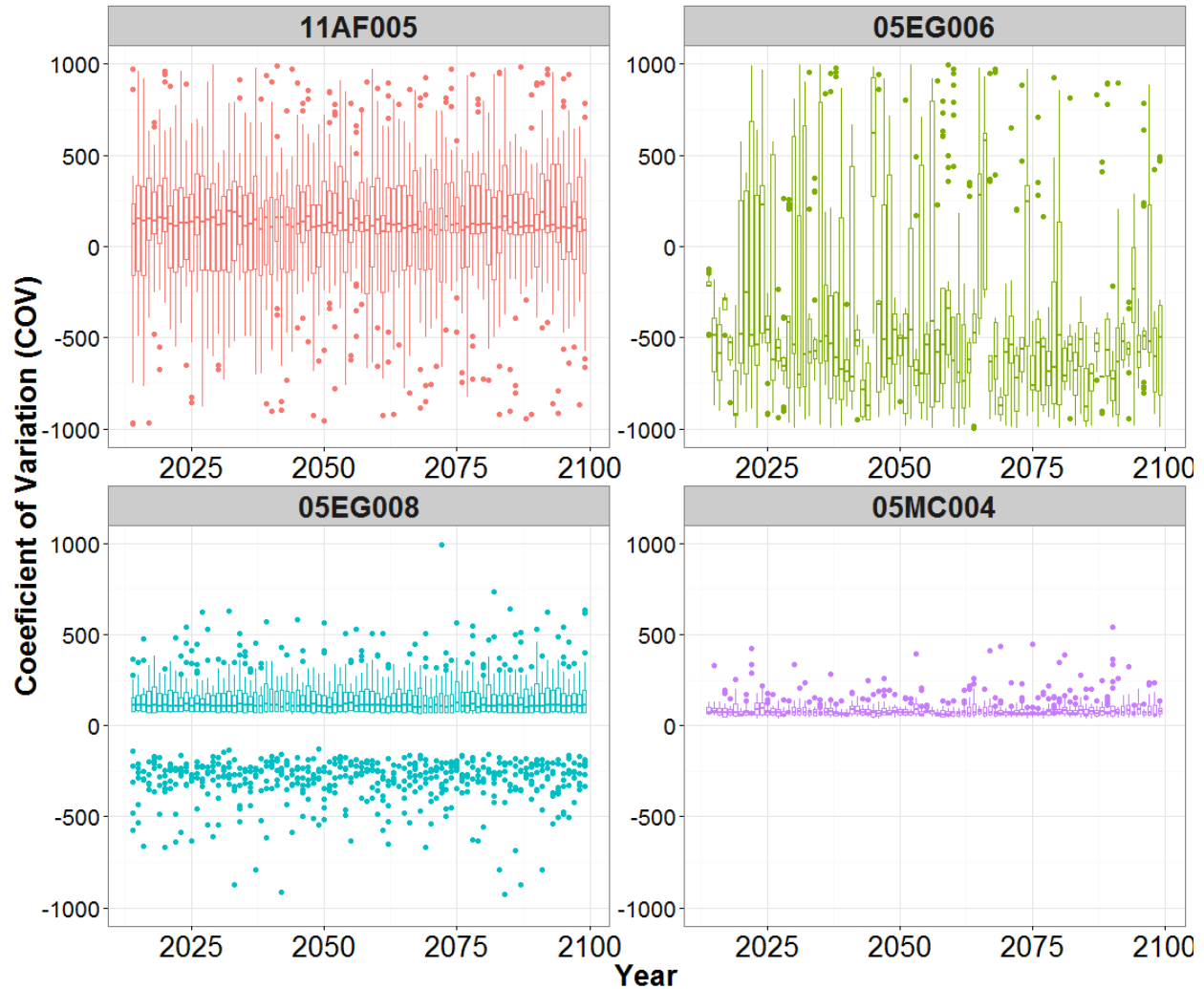
Figure 5.12. Annual boxplots showing temporal variation in PSC values for streamflow at catchment outlet.

It is found that the PSC magnitudes are dependent on the choice of GCM, RCP and downscaling model. Figure 5.13 shows the sensitivity of PSC values towards above mentioned sources of uncertainty for precipitation, temperature and flow at all catchments. It is found that the magnitudes of PSC are most sensitive to the choice of downscaling model (D.mod in Figure 5.13) in case of flow and precipitation. Among different downscaling models considered, largest PSC magnitudes for precipitation are obtained in catchments

05MC004 and 11AF005 when SPS5x5 is considered as the downscaling model. SPS7x7 model brings most uncertainty in catchment EG008 while SPS3x3 model is found to produce largest PSC magnitudes in the 05EG006 catchment. In case of streamflow, SPS7x7 model brings the largest uncertainty in the catchment 05EG008 while model SPS5x5 model brings largest uncertainty in the catchments: 05MC004 and 11AF005. Model SP is found to be associated with lowest PSC values in the all the scenarios analyzed. The choice of GCM is found to be most critical for temperature related PSCs as seen in Figure 5.13. Among the three GCMs analyzed temperature PSC magnitudes in catchment 11AF005 are found to be most sensitive to projections made by the GCM: iap_fgoals while at other catchments PSC magnitudes are found to be most sensitive to the GCM: mri_cgcm3.
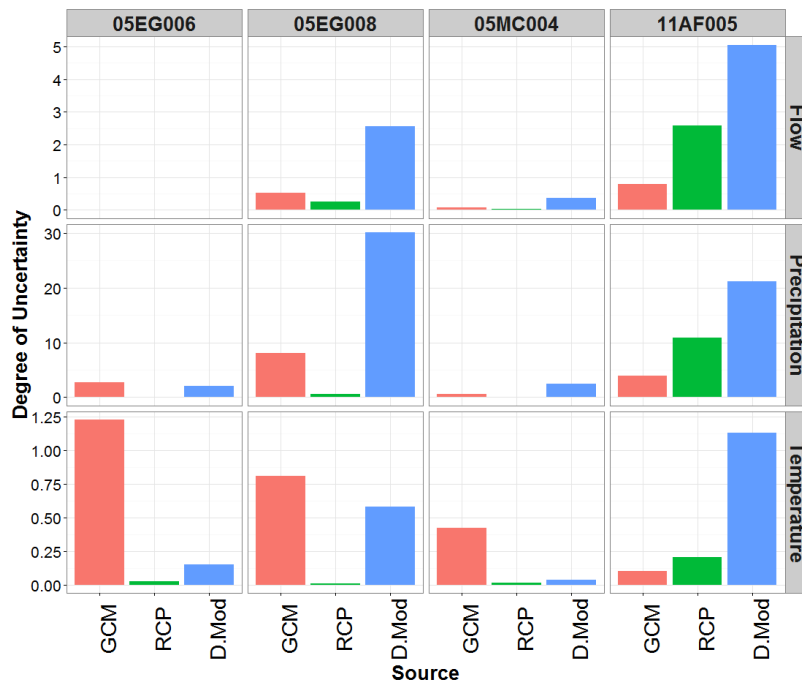

Figure 5.13. Comparison of uncertainty sources influencing PSC magnitudes for catchment averaged precipitation, temperature and streamflow. The degree of uncertainty is expressed in

terms of percentage for precipitation and flows while absolute changes are provided for temperature. Note that catchment 05EG006 doesn't show any significant changes in streamflow.

Lastly, the inclusion of catchment characteristics into the downscaling process is also found to influence flood magnitudes. Among the different sources of uncertainty, changes in flooding magnitudes are found to be most significantly influenced by the return period of flood being analyzed and the choice of the downscaling method as illustrated in Figure 5.14.A summary of the results is presented in Figure 5.15 for all return periods and downscaling models considered. The results presented include projections made by all GCM-RCP combinations. It can be noticed that larger increases in flood magnitudes are obtained for low return period events than for high return period events. An average increase of 3%, 1%, 0.7%, 0.4% and 0.3% in flood magnitudes is obtained for return periods 2, 5, 10, 25 and 100 year return periods respectively. The largest PSC magnitudes in flood extremes are obtained for the catchment 05EG008 (4%), followed by catchment 05MC004 (1%), followed by 05EG006 (-0.1%), and followed by catchment 11AF005 (-0.5%). Further largest increases in flood extremes are obtained considering SPS5x5 as the test model (3%), followed by SPS7x7 model (2%), followed by SPS3x3 model (1%), followed by SP model (0.2%), and followed by SPS9x9 model (-0.4%).
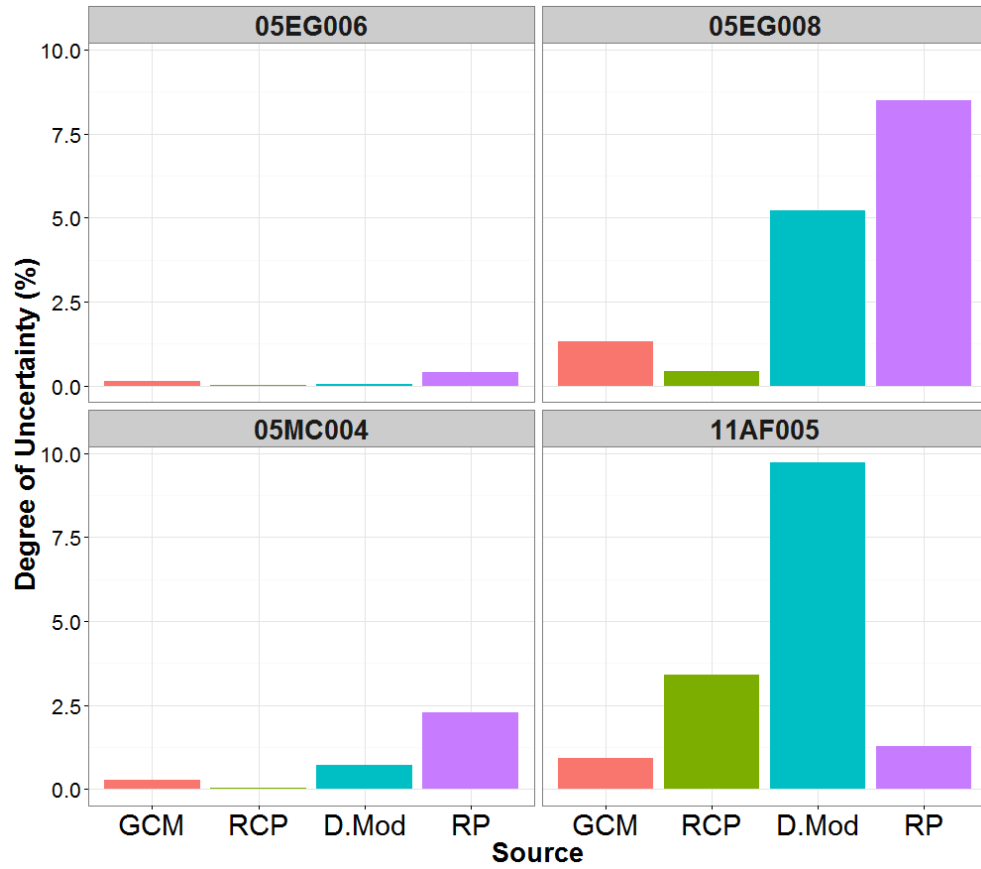
Figure 5.14. Comparison of uncertainty sources influencing PSC magnitudes for flood events.
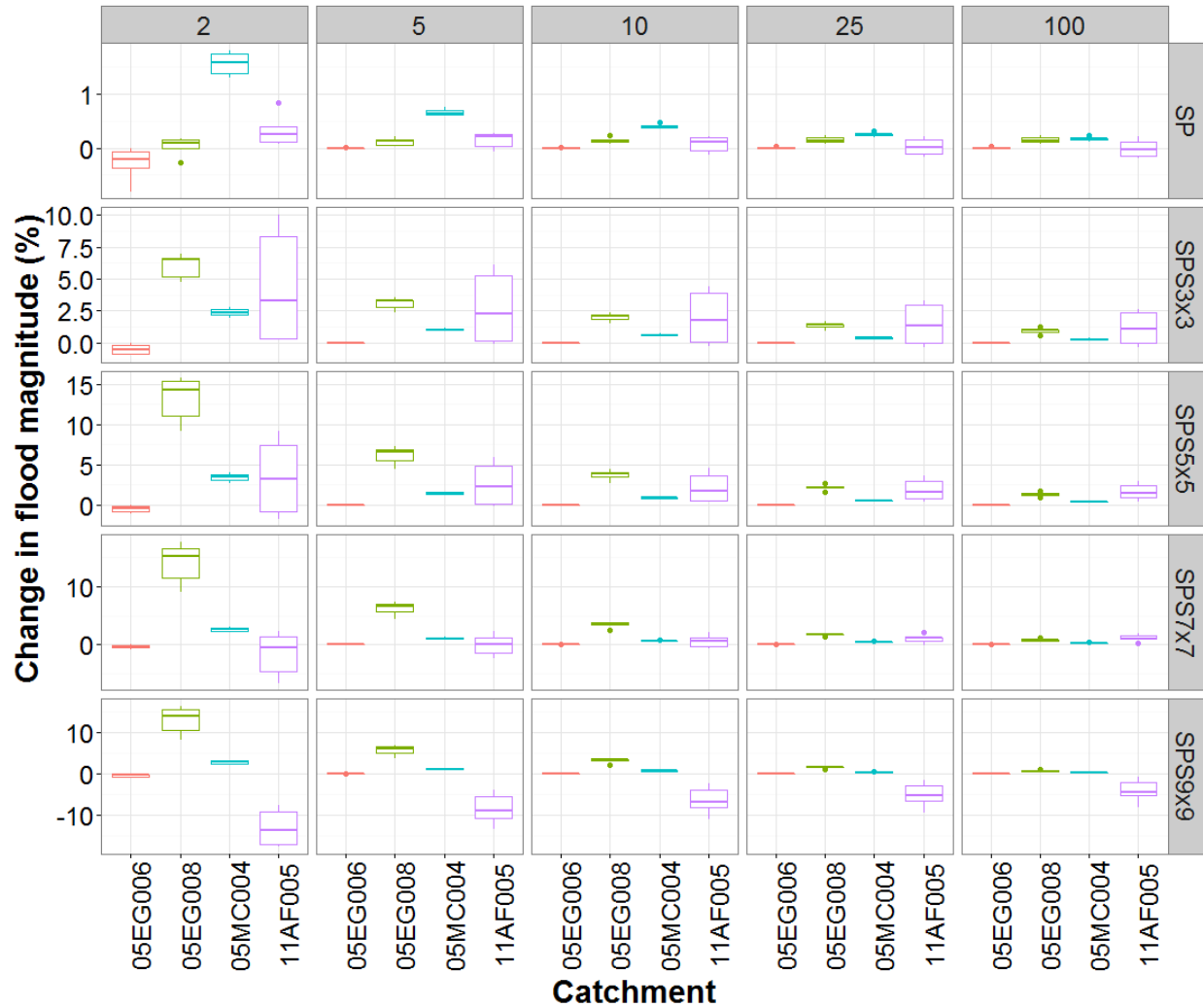
Figure 5.15. PSC magnitudes for floods of different intensities at all catchments analyzed.

5.7 Conclusions

In this study the climatic impacts of catchment physical characteristics are quantified and their hydrologic implications are explored. Four small catchments located in southern Saskatchewan region are selected for study. Future climate and hydrologic projections are made by downscaling GCM outputs using different SP and SPS method based models. The

downscaled climate projections are used with Sacramento hydrologic model to derive future streamflow response from all four catchments. The climatic and hydrologic response derived from projections downscaled using model: SP_LC_elev is compared with the projections obtained from downscaling models: SP, SPS3x3, SPS5x5, SPS7x7 and SPS9x9 and the differences between them are explored. Further similar analysis is also performed with the peak flows.

From this analysis certain key conclusions can be drawn:

- The SP and SPS based models can be used to project climatic and hydrologic response of a catchment taking into consideration their physical characteristics.

- Physical characteristics of a catchment do influence its climatic regime which in turn has hydrologic implications. These influences are found to be statistically significant in approximately 70% of the scenarios analyzed. Precipitation is found to be the most influenced variable, followed by catchment outflow while temperature is found to be the least influenced.

- Huge spatial and temporal variability in PSC magnitudes across the catchment area and over the projection period is noted in this study for precipitation, temperature and flow variables. The spatial variability in flow related PSC magnitudes can be analyzed using a distributed hydrologic model, which can be one possible area of extension of the current study.

- A comparison of uncertainty sources influencing precipitation, temperature and flow related PSC magnitudes shows that the changes are most significantly influenced by the choice of the downscaling model for precipitation and flow while the choice of GCM dominates the temperature related PSC magnitudes.

- SPS method based models in general are found to have higher PSC magnitudes than SP method based model.

- PSC in climatic variables are also found to influence hydrological extremes. It is found that the changes are higher for low return period events than the high return period events which suggest that small scale more frequent flooding events are more impacted by physically induced climatic changes than the large scale rare flooding events. Among the different sources of uncertainty influencing changes in hydrological extremes, the choice of downscaling method as well as return period of the flooding event being analyzed are found to be the two most important factors influencing PSC magnitudes.

Above results establish the existence of significant climatic and hydrologic PSCs at a catchment scale and highlight the importance of considering them while making future flow projections.

The current study aimed at making the first step towards quantification of physically sourced climatic and hydrologic changes within a statistical downscaling framework. It will be interesting to extend this analysis to other catchments which are larger than the ones selected in this study as well as are located in different biomes and compare and contrast the results to have a better understanding of physically sourced climatic and hydrologic changes.

References

Burnash, RJC (1995) The NWS River Forecast System – Catchment Modeling. In: Vijay P. Singh (ed.), Computer models of watershed hydrology. Revised edition, Highlands Ranch, Colo.: Water Resources Publications, c1995. http://www.wrpllc.com/books/cmwh.html.

Chiew FHS, Kirono DGC, Kent DM, Frost AJ, Charles SP, Timbal B, Nguyen KC, Fu G (2010) Comparison of runoff modelled using rainfall from different downscaling methods for historical and future climates. Journal of Hydrology 387: 10-23.

Coles S (2001) An Introduction to Statistical Modeling of Extreme Values (Springer, London).

Duan Q, Gupta VK, Sorooshian S (1993) A shuffled complex evolution approach for effective and efficient optimization. J. Optimization Theory Appl., 76(3): 501-521.

Dwarakish GS, Ganasri BP (2015) Impact of land use change on hydrological systems: A review of current modeling approaches. Cogent Geoscience 1:1115691.

Encyclopedia of Saskatchewan (2016) Climate. University of Regina, Accessed March 7 2016. [Available online at http://esask.uregina.ca/entry/climate.html].

Flato G, Marotzke J, Abiodun B, et al. (2013) Evaluation of Climate Models. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker TF, Qin D, Plattner GK, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V and Midgley PM (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Frost AJ, Charles SP, Timbal B, Chiew FHS, Mehrotra R, Nguyen KC, Chandler RE, McGregor JL, Fu G, Kirono DGC, Fernandez E, Kent DM (2011) A comparison of multi-site daily

rainfall downscaling techniques under Australian conditions. Journal of Hydrology 408: 1-18.

Gaur A, Simonovic SP (2016a) A Scaling Method for Physically Representative Downscaling of Climate Model Data. Under review in Climate Dynamics.

Gaur A, Simonovic SP (2016b) Extension of SP method and its application towards downscaling climate model based near surface air temperature. Under review in Journal of Applied Meteorology and Climatology.

Gaur A, Simonovic SP (2016c) A Scaling Method for Physically Representative Downscaling of Climate Model Data. Under review in Theoretical and Applied Climatology.

Gaur and Simonovic (2015) Projected Changes in the Dynamics of Flood Hazard in the Grand River Basin, Canada. British Journal of Environment and Climate Change 5(1): 37-51.

Grillakis MG, Koutroulis AG, Tsanis IK (2011) Climate change impact on the hydrology of spencer creek watershed in Southern Ontario, Canada. J Hydrol. 409: 1-19.

Hopkinson RF et al. (2011) Impact of aligning climatological day on gridding daily maximum-minimum temperature and precipitation over Canada. Journal of Applied Meteorology and Climatology 50: 1654:1665.

Hosking JRM (1990) L-Moments: Analysis and Estimation of Distributions Using Linear Combinations of Order Statistics. Journal of the Royal Statistical Society (Series B) 52: 105-124.

Hurtt GC, et al. (2011) Harmonization of land-use scenarios for the period 1500–2100: 600 years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands. Clim. Change 109(1-2): 117–161.

Hutchinson MF et al. (2009) Development and testing of Canada-wide interpolated spatial models of daily minimum-maximum temperature and precipitation for 1961-2003. Journal of Applied Meteorology and Climatology 48: 725:741.

Jarvis A, Reuter HI, Nelson A, Guevara E (2008) Hole-filled SRTM for the globe Version 4, available from the CGIAR-CSI SRTM 90m Database (http://srtm.csi.cgiar.org).

Kingston DG, Taylor RG (2010) Sources of uncertainty in climate change impacts on river discharge and groundwater in a headwater catchment of the Upper Nile Basin, Uganda. Hydrol Earth Syst Sc. 14(7):1297-1308.

King LM, McLeod AI, Simonovic SP (2015) Improved Weather Generator Algorithm for Multisite Simulation of Precipitation and Temperature. Journal of the American Water Resources Association (JAWRA) 51(5): 1305-1320.

Leander R, Buishand TA (2007) Resampling of regional climate model output for the simulation of extreme river flows. J Hydr. 332(3-4):487–496.

Mesinger F, DiMego G, Kalnay E, Mitchell K, and Coauthors (2006) North American Regional Reanalysis. Bulletin of the American Meteorological Society 87: 343–360.

Nash JE, Sutcliffe JV (1970) River flow forecasting through conceptual models, Part I - A discussion of principles. J. Hydrol. 10: 282–290.

Nelder JA, Mead R (1965) A simplex method for function minimization, Computer Journal 7: 308–313.

Schneider C, Laizé CLR, Acreman MC, Flörke M (2012) How Will Climate Change Modify River Flow Regimes in Europe? Hydrol Earth Syst Sc. Dis.9(8): 9193–9238.

Schoof JT (2013) Statistical downscaling in climatology. Geogr. Compass 7: 249–265.

Srivastav R, Simonovic SP (2014) Multi-site, multivariate weather generator using maximum entropy bootstrap. Clim Dyn. 44 (11): 3431-3448.

Taylor KE, Stouffer RJ, Meehl GA (2012) An overview of CMIP5 and the experimental design. Bulletin of the American Meteorological Society 93: 485–498.

Te Linde AH, Aerts JCJH, Bakker AMR, Kwadijk JCJ (2010) Simulating low-probability peak discharges for the Rhine basin using resampled climate modeling data. Water Resour. Res. 46: W03512.

Tejeda EM, Zabalza J, Rahman K, Gago-Silva A, Lopez-Moreno JI, Vincente-Serrano S, Lehmann A, Tague C, Beniston M (2014) Hydrological impacts of climate and land-use changes in a mountain watershed: uncertainty estimation based on model comparison. Ecohydrol. 8(8): 1396-1416.

Thanapakpawin P, Richey J, Thomas D, Rodda S, Campbell B, Logsdon M (2006) Effects of landuse change on the hydrologic regime of the Mae Chaem river basin, NW Thailand. Journal of Hydrology 334: 215-230.

Verburg PH, de Nijs TCM, van Eck JR, Visser H, de Jong K (2004) A method to analyse neighbourhood characteristics of land use patterns. Computers, Environment and Urban Systems 28: 667–690.

Walter MT, Brooks ES, McCool DK, King LG, Molnau M, Boll J (2005) Process-based snowmelt modeling: Does it require more input data than temperature-index modeling? Journal of Hydrology 300(1-4): 65-75.

Wilby RL, Dawson CW, Barrow EM (2002) SDSM – a decision support tool for the assessment of regional climate change impacts. Environmental Modelling & Software 17(2): 145–157.

Wilcoxon F (1945) Individual comparisons by ranking methods. Biometrics Bulletin 1 (6): 80–83.

Wood AW, Leung LR, Sridhar V, Lettenmaier DP (2004) Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. Climatic Change 62: 189–216.

Wood SN (2000) Modelling and Smoothing Parameter Estimation with Multiple Quadratic Penalties. J.R. Statist. Soc. B 62(2):413-428.

Zorita E, von Storch H (1999) The analog method as a simple statistical downscaling technique: comparison with more complicated methods. Journal of Climate 12: 2474–2489.

CHAPTER 6: Conclusions

In this section major conclusions from this thesis are summarized and ideas for further research in this direction are provided.

6.1 Summary and conclusions

The research presented in this thesis revolves around the quantification of future climatic and hydrologic changes that are physically sourced. Earlier studies have performed similar analysis but within a dynamic downscaling framework. In this study a statistical downscaling based methodology has been used.

To achieve this objective, a Physical Scaling (SP) method of downscaling GCM projections is developed. The proposed model formulation considers land-cover, elevation and their distribution across the study region as predictors in addition to GCM data. The inclusion of physical parameters in the model formulation provides the Physical Scaling method with an additional capability of exploring the effects of changes in physical characteristics on local climate. The proposed Physical Scaling method is tested for its ability to downscale important climatic variables like precipitation, surface and air temperatures across the southern Saskatchewan region of Canada. It is found that the model is capable of downscaling above mentioned climatic variables across this region accurately and hence is used to identify physically sourced climatic and hydrologic changes in four small catchments located across this region.

A comparison of downscaled temperature, precipitation and flow variables obtained with and without considering catchment physical characteristics indicate statistically significant differences between them in nearly 70% of the cases analyzed. It is found that the differences

are more pronounced in case of precipitation and flow than in the case of temperature. A comparison of the magnitude of physically sourced changes for different versions of SP and SPS method based models indicate that SPS method based models are associated with larger changes than the SP method based models. The downscaled precipitation, temperature and flow magnitudes are found to vary with the choice of GCM, RCP, and the version of SP and SPS method based downscaling model. Among these sources of uncertainty, the choice of downscaling model version is found to impact the magnitude of physically sourced changes most significantly. In the case of flood magnitudes, the choices of return period of flood event as well as the downscaling model version are found to be the two most important factors governing the magnitudes of physically sourced changes. Further large spatial and temporal variations in physically sourced changes are obtained. The results obtained in this study highlight that changes owing to changes in physical characteristics of a region can be significant at a local scale and hence should be accounted for while making future hydrologic predictions. This step makes a first step towards doing it within a statistical downscaling framework.

The Physical Scaling method introduced in this study has following limitations. The spatial resolution of the downscaled output obtained from this method is limited by the spatial resolution of the land-cover data. In this thesis MODIS land-cover data have been used that have a spatial resolution of 500 m and hence the spatial resolution of the downscaled outputs is also 500 m. In future with the development of continuous, high resolution remotely sensed data this limitation can be overcome and even higher resolution downscaled products can be obtained. Further model formulation is based on the assumption that the effects of large scale climate, elevation, land-cover and neighborhood configuration are additive in nature. In future several

dimensionality reduction methods like Principal Component Analysis (PCA) can be explored while linking predictors with predictant. Further usage of machine learning and cognitive science based methods like Artificial Neural Network (ANN) can be explored. Lastly in this study three GCMs have been used to make future climatic and hydrologic projections however other models can be used in future and GCM related uncertainty can be explored.

6.2 Areas of further research

The presented research can be extended in many possible directions. These include:

- SP method can be modified to account for snow-cover during the winter months. Snow-cover is a very important physical parameter which effects the climatology and hydrology of any region. It is especially relevant in Canadian conditions and can be included into SP method definition in future.

- An appropriate spatial scale for the calibration of SP method needs to be ascertained. The spatial scale chosen should be a compromise between accuracy and robustness of the downscaled results.

- The applicability of SP method can be evaluated in other regions of Canada and the globe. It will be interesting to see how the model performs in regions that have more complex physiography than the region considered in this study. Further, case studies can be performed on catchments located in other regions of the globe to better understand the underlying dynamics of the physically driven changes.

- It will be interesting to compare the direct land-cover driven hydrologic changes with the indirect physically induced hydrologic changes of land-cover. The study can be

performed at catchments located in different climatic regions and biomes to compare and contrast the results.

- The methodology developed for land-cover downscaling in this thesis can be used to downscale land-use projections for multiple GCMs and scenarios to compare land-cover projections across the globe.

Appendices

Appendix A. Land-cover based variation in surface temperature in the southern Saskatchewan region over the period 2006-2013.

Figure below shows the variation of temperature observed over the period 2006-2013 for different land-cover classes present within the study region. Large variations in temperature owing to large scale climate are also evident from the figure.
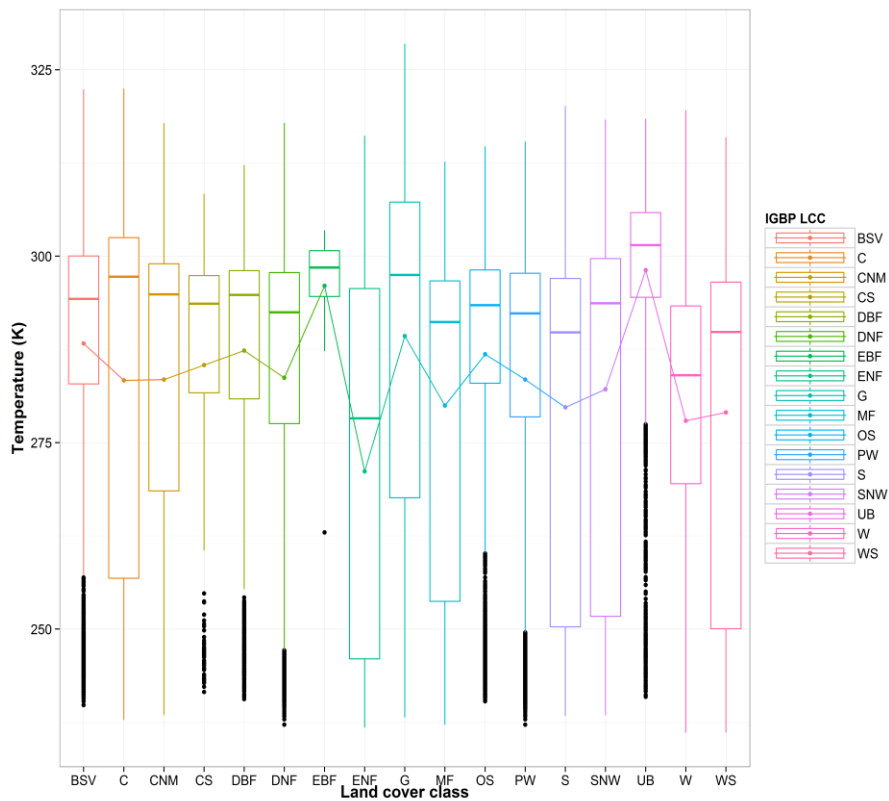

Figure A1. Mean land-cover specific surface temperatures across the study region.

Appendix B. Spatial distribution of model error over the southern Saskatchewan region.

Figure below provides spatial distribution of model error and spatial distribution of elevation across the study region. It can be seen that the model performs better in low elevation regions than the high elevation regions.
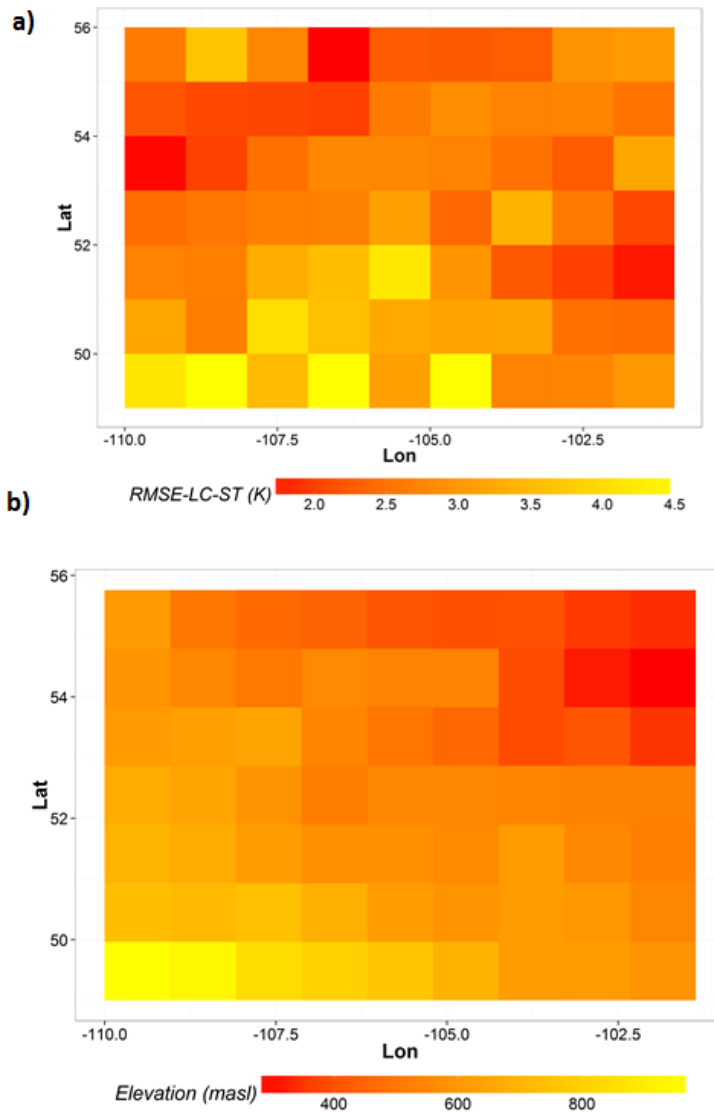

Figure B1. a) Spatial distribution of the RMSE$_{st-lc}$ statistic across the study region, **b)** variation in elevation across the study region.

Appendix C. Increase in spatial resolution in the downscaled air temperature as obtained from the indirect approach.

Using the indirect approach to downscaling GCM based air temperature, a regular mesh of downscaled air temperature can be obtained. Following figure shows a sample result obtained for GCM: iap-fgoals and RCP2.6 for the year 2100 using SPS3x3, SPS5x5, SPS7x7 and SPS9x9 downscaling models.
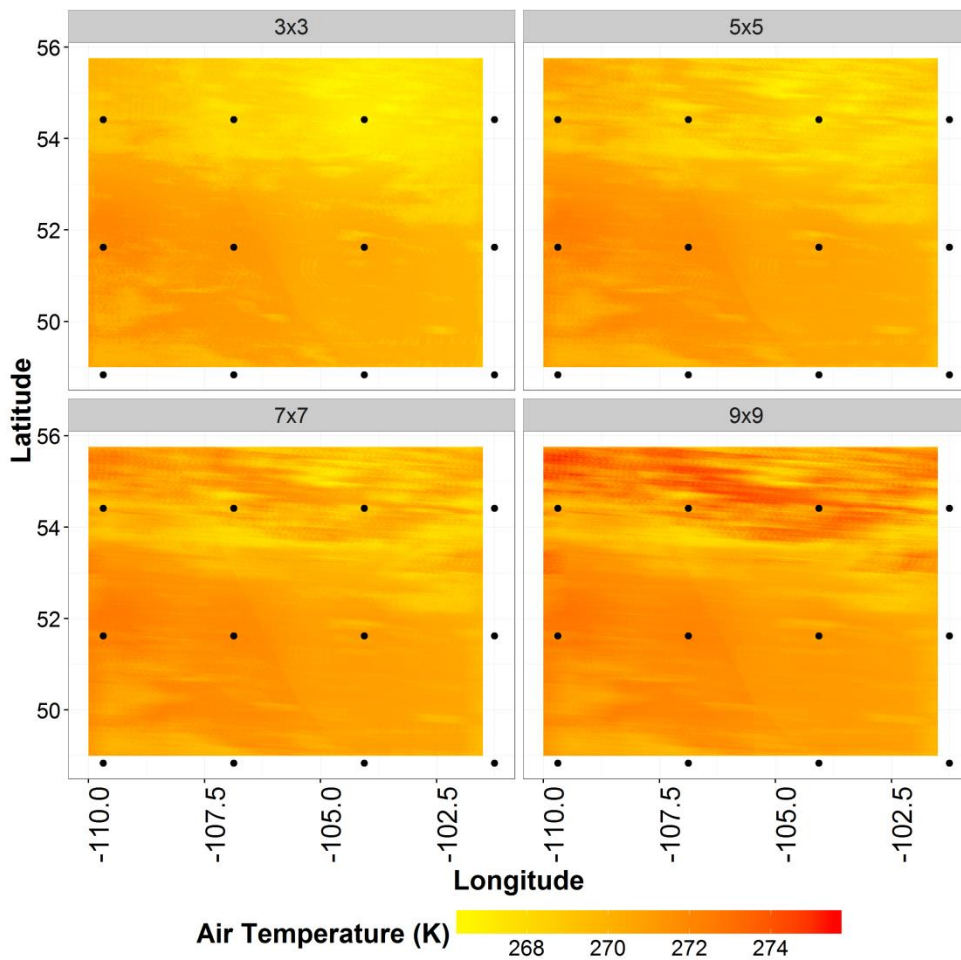

Figure C1. Increase in spatial resolution in the downscaled air temperature as obtained from the indirect approach.

Appendix D. Sacramento model parameters and parameter range considered in this analysis.

The Sacramento hydrologic model chosen for analysis in this study has 13 free parameters (H1–H13) and the routing scheme considered has three free parameters (R1-R3). A list of these parameters, their description and the valid range of values is provided below.

Table D1. Sacramento model parameters and parameter range considered in this analysis.

S.No.	Parameter	Description	Range
H1	uztwm	The upper layer tension water capacity, mm	10-300
H2	uzfwm	The upper layer free water capacity, mm	5-150
H3	uzk	Interflow depletion rate from the upper layer free water storage, day^{-1}	0.10-0.75
H4	pctim	Permanent impervious area fraction	0-1
H5	adimp	Maximum fraction of an additional impervious area due to saturation	0-1
H6	zperc	Ratio of maximum and minimum percolation rates	5-350
H7	rexp	Shape parameter of the percolation curve	1-5
H8	lztwm	The lower layer tension water capacity, mm	10-500
H9	lzfsm	The lower layer supplemental free water capacity, mm	5-400
H10	lzfpm	The lower layer primary free water capacity, mm	10-1000
H11	lzsk	Depletion rate of the lower layer supplemental free water storage, day^{-1}	0.01-0.35
H12	lzpk	Depletion rate of the lower layer primary free water storage, day^{-1}	0.001-0.05
H13	pfree	Percolation fraction that goes directly to the lower layer free water	0-0.8
R1	tau_s	time constant for the slow flow component	5-100
R2	tau_q	time constant for the quick flow component	0-5
R3	v_s	fractional volume for the exponential component	0-1

CURRICULUM VITAE

EDUCATION

2013-current London, Ontario	*Doctor of Philosophy (PhD)* Western University
2011-2013 London, Ontario	*Master in Engineering Science (MESc)* Western University
2006-2010 Durgapur, India	*Bachelor of Technology (BTech)* National Institute of Technology, Durgapur

RESEARCH EXPERIENCE

May 2011-current	Western University (Canada)
May 2010 – December 2010	Indian Institute of Technology, Roorkee (India)
May 2009 - April 2010	National Institute of Technology, Durgapur (India)
May-August 2009	Technical University of Hamburg (Germany) *German Academic Exchange Service (DAAD)*

TEACHING EXPERIENCE

May 2011-current	*Teaching assistant* Western University (Canada)
April 20-22, 2015	*Instructor* Academia Sinica (Taiwan)

VOLUNTEERING EXPERIENCE

2013-current	*Coordinator* Senior Youth Group, Hindu Cultural Center, Canada
2015-current	*Treasurer* Western Water Chapter, Western University, Canada
2007-2009	*Coordinator* Kartavya (NGO) Durgapur, India
2007-2010	*Student Representative* Civil Engineering department, NIT Durgapur, India

JOURNAL PUBLICATIONS

Gaur A, Simonovic SP (2015) Towards Reducing Climate Change Impact Assessment Process

Uncertainty. Environmental Processes 2(2): 275-290.

Gaur A, Simonovic SP (2015) Projected Changes in the Dynamics of Flood Hazard in the Grand

River Basin, Canada. British Journal of Environment and Climate Change 5(1): 37-51.

CONFERENCE PROCEEDINGS

Gaur A, Simonovic SP. Potential Use of an Open-Source Software R as a Tool for Performing

Climate Change Impact Studies. ed. Proceedings of DailyMeteo.org/2014 Conference.

Belgrade, Serbia 26-27 June 2014.

Kansal M, Gaur A (2011) Expert System Based Water Sustainability Index. World

Environmental and Water Resources Congress 2011: pp. 1692-1704.

Bhattacharya K, Gaur A, Chakraborty I, Measured and Modelled Response of a Masonry Wall

under out-of-plane Excitation. 14th Symposium on Earthquake Engineering, IIT Roorkee.

17-19 December 2010.