

Probabilistic Event Attribution of Precipitation and Streamflow Extremes in the Spring 2022 Manitoba Flood Using LSTM-Based Hydrological Modeling

By

Muhammad Zain Ul Abideen

A Thesis submitted to the Faculty of Graduate and Postdoctoral Studies in
partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Department of Civil Engineering

University of Manitoba

Winnipeg, Manitoba, Canada

August 2026

Copyright © 2026 by Muhammad Zain Ul Abideen

Abstract

The Spring 2022 flood was among the most extensive and damaging on record in southern Manitoba, yet the contribution of anthropogenic climate change to such cold-region events remains poorly quantified. This thesis quantifies the influence of anthropogenic climate change on the Spring 2022 precipitation and flood event using a probabilistic event attribution framework applied across both meteorological and hydrological domains. The observational records from 19 meteorological stations and 28 HYDAT streamflow gauges spanning the period 1950–2023 were considered, with the Whitemouth River sub-basin (gauge 05PH003) serving as the focal catchment. Long Short-Term Memory (LSTM) networks reconstructed six incomplete streamflow records. The meteorological event was the most extreme 29-day April–May precipitation accumulation, and the hydrological event the most extreme 27-day mean streamflow. Attribution used three approaches contrasting factual and counterfactual climates: General Circulation Model (GCM)-based, non-stationary scaling, and time-slice methods. Streamflow was simulated with LSTM networks and the Raven process-based model forced by climate model inputs. To explicitly account for the role of antecedent snow conditions in flood generation, a compound snow–temperature preconditioning analysis was conducted at Indian Bay using bivariate copulas, jointly characterising the influence of winter snowpack accumulation and early spring temperatures as compound precursors to flooding. Median risk ratios (RR) indicated that precipitation magnitude

was likely exacerbated by anthropogenic forcing (1.00 to 1.53), with the strongest signal in southeastern Manitoba, underlain by the Canadian Shield and behaves distinctly from the lacustrine Red River Valley to the west. Flood attribution did not reach statistical significance: streamflow risk ratios leaned above unity but were not robust across either approach (LSTM: RR 1.45 [0.65, 2.66]; Raven: RR 3.14 [0.83, 8.48]), and the change in compound snow–temperature preconditioning was non-significant (RR 1.11 [0.13, 5.12]). These results reveal a partial decoupling between precipitation and flood attribution in cold-region Prairie basins, where flood magnitude is governed by snowpack, melt timing, antecedent wetness, and storage rather than precipitation intensity alone, so precipitation-based attribution cannot serve as a flood-risk proxy. The findings advance climate change attribution in snow-dominated basins and can inform flood risk assessment and adaptation planning across the Canadian Prairies.

Acknowledgments

First and foremost, I would like to thank God for granting me the strength, patience, and perseverance to reach this stage of my journey.

I am deeply grateful to my supervisor, Assistant Professor Dr. Chandra Rajulapati, for her constant support, insightful advice, and invaluable mentorship throughout the course of this research project.

I am also thankful to the faculty and staff of the Department of Civil Engineering at the University of Manitoba for providing an enriching academic environment and the resources necessary to carry out this research. Special thanks to my workspace mates and colleagues for their collaboration, technical assistance, and encouragement along the way.

I would like to express my gratitude to the committee members, Dr. Alex Crawford and Dr. Ricardo Mantilla, for providing valuable comments.

I would like to express my appreciation to the funding organizations, including Mitacs and Manitoba Hydro. Your monetary support has made this research possible.

Finally, I want to express my heartfelt appreciation to my family and friends for their unwavering support, understanding, and belief in me. Their encouragement and love have been a constant source of strength throughout every stage of this journey.

Table of Contents

Abstract	II
Acknowledgments.....	IV
Table of Contents	V
List of Tables	X
List of Figures.....	XI
Chapter 1 Introduction	1
1.1 Background and Motivation	1
1.1.1 Flood Risk in Cold-Region Basins	1
1.1.2 Prairie Hydrology and Snow-Dominated Systems.....	2
1.1.3 Anthropogenic Climate Change and Extreme Precipitation.....	3
1.1.4 Challenges in Flood Attribution.....	4
1.2 Research Objectives	6
1.3 Thesis Structure	7
Chapter 2 Study Area, Data, and Preprocessing Framework.....	8
2.1 Study Area.....	8

2.2 Data	11
2.2.1 Streamflow Observations (HYDAT).....	11
2.2.2 Precipitation and Temperature Observations.....	12
2.2.3 ERA5 Reanalysis.....	15
2.2.4 Snow Depth Observations	15
2.2.5 CMIP6 Climate Model Simulations	16
2.3 Data Preprocessing and Quality Control.....	18
2.3.1 Dataset Screening and Quality Control	18
2.3.2 Spatial Processing.....	19
2.4 Event Definition and Extreme Value Framework	19
2.4.1 Event Definition	20
2.4.2 Extreme Value Theory.....	20
Chapter 3 LSTM-Based Streamflow Gap-Filling Framework.....	21
3.1 Motivation for Gap Filling	21
3.2 LSTM Modeling Framework.....	22
3.2.1 Thiessen-Based Forcing Configuration	22

3.2.2 ERA5-Based Forcing Configuration	25
3.2.3 Comparison Design	25
3.3 Gap-Filling Implementation	26
3.4 Results: Gap-Filling Performance	30
3.4.1 Station-Wise Performance Comparison	31
3.4.2 Forcing Selection.....	32
3.5 Implications for Subsequent Analyses	34
Chapter 4 Assessing the Role of Climate Change in the 2022 Manitoba Spring Flood.....	37
4.1 Introduction	38
4.2 Study Area and Data	40
4.3 Methods	41
4.3.1 Event Definition and Rarity	41
4.3.2 Attribution Methods	44
4.3.3 Bivariate (Compound) Attribution Framework (Snow-Temperature)	46
4.3.4 Hydrological Modeling.....	49
4.4 Results	51

4.4.1 Scaling-Based PEA Analysis	51
4.4.2 Time-Slice PEA Analysis	55
4.4.3 GCM-Based PEA Analysis for Precipitation.....	60
4.4.4 Performance of the LSTM and Raven Streamflow Simulation	65
4.4.5 GCM-Based PEA Analysis for Streamflow	67
4.4.6 Bivariate (Compound) Attribution Results.....	70
4.5 Discussion and Conclusions	74
4.5.1 Diagnostic Synthesis of the Regional Attribution Signal	75
4.5.2 Hydroclimatological Conditions for the 2022 Flood	78
4.5.3 Broader Implications and Future Directions.....	80
4.6 Data Availability.....	81
Chapter 5 Conclusions and Future Directions.....	83
5.1 General Summary	83
5.2 Recommendations for Future Research	85
References	87
Appendix A: Supplementary Tables and Figures	101

Appendix B: LSTM Training and Hyperparameter Diagnostics	106
Appendix C: Additional EVT and Attribution Diagnostics	111
Appendix D: Additional Copula Diagnostics	119

List of Tables

Table 1: Summary of selected HYDAT streamflow gauges used in this study	13
Table 2: Summary of meteorological, reanalysis, snow-related, and climate-model datasets used in this thesis.....	17
Table 3: Architecture and training hyperparameters of the regional LSTM models under the Thiessen and ERA5 forcing configurations.	23
Table 4: Dynamic meteorological inputs and static basin attributes used in the Thiessen and ERA5 LSTM configurations.....	24
Table 5: Summary of the six-target gap-filled stations	26
Table A1: CMIP6 models used in this study, including nominal horizontal resolution, simulation periods, and modeling centers.....	101
Table A2: Schematic mapping of research objectives to the corresponding variables, methodological approaches, and data sources used in this study.....	102
Table D1: Best-fitting marginal distributions and copula families by station and period, from the exploratory multi-station screening.....	119

List of Figures

Figure 1: Study area and hydroclimatic context for the 2022 spring flood in southern Manitoba.	10
Figure 2: Gap-filling diagnostic for the six target stations under the Thiessen-forcing regional LSTM.....	28
Figure 3: As Figure 2, but under the ERA5-forced regional LSTM adopted for the final gap-filled product.	29
Figure 4: Epoch-wise validation performance of the two regional gap-filling LSTMs.....	30
Figure 5: Station-wise comparison of ERA5 and Thiessen meteorological forcings for the regional LSTM across the seven study sub-basins	33
Figure 6: Empirical cumulative distribution functions (ECDFs) of annual maximum RXnD streamflow before (red) and after (blue) LSTM-based gap filling for the six gap-filling target stations.....	36
Figure 7: Event-based heatmaps of return periods estimated using the generalized extreme value (GEV) distribution for April to May 2022.....	42
Figure 8: Spatial distribution of station-level risk ratios (RR) derived from the scaling-based attribution method.	53

Figure 9: Return-level curves and corresponding risk ratios (RR) for annual maximum precipitation at 49.02°N, 95.60°W	55
Figure 10: Spatial distribution of station-level risk ratios (RR) derived from the time-slice attribution method.	58
Figure 11: Return-level curves and corresponding risk ratios (RR) for annual maximum 29-day precipitation at 50.52°N, 95.98°W	60
Figure 12: Spatial distribution of station-level risk ratios (RR) derived from the GCM-based attribution method, using the four-model CMIP6 ensemble mean.	63
Figure 13: Return-level curves and corresponding risk ratio (RR) for annual maximum 29-day precipitation at the representative station at 49.83°N, 92.80°W	64
Figure 14: Observed (blue) and simulated (orange) daily streamflow at the Whitemouth River gauge (05PH003).....	67
Figure 15: Ensemble-mean return-level curves and corresponding risk ratios (RR) for annual maximum 27-day mean flow at the Whitemouth River gauge (49.94°N, 95.96°W) from the GCM-based attribution method.....	70
Figure 16: Bivariate compound attribution results for cool-spring and high-snow preconditioning at Indian Bay (5031320).	73

Figure A1: Daily precipitation climatology over the study region for the 1950–2020 period, comparing observed station data (black) with historical simulations from four CMIP6 models.....	104
Figure A2: Schematic overview of the probabilistic event attribution framework applied in this study.....	105
Figure B1: Validation performance across training epochs: (a) median Nash-Sutcliffe Efficiency (NSE) and Kling-Gupta Efficiency (KGE), and (b) median Root Mean Square Error (RMSE).....	106
Figure B2: Flow duration curves (FDCs) for station 05PH003 comparing LSTM (a–c) and Raven (d–f) simulated daily streamflow against observations on a logarithmic discharge axis.	107
Figure B3: Hyperparameter grid search for the Thiessen-forcing regional LSTM: 81 configurations spanning hidden size, batch size, output dropout, and sequence length.....	108
Figure B4: As Figure B3, but for the ERA5-forcing experiment (sequence lengths of 180, 270, and 365 days).....	109
Figure B5: As Figure 6, but for the Thiessen-forced reconstructions; shown for completeness of the forcing comparison	110

Figure C1: GEV diagnostics for 29-day maximum precipitation (RX29D) at Indian Bay.....	111
Figure C2: GEV diagnostics for 27-day mean streamflow at the Whitemouth River sub-basin (05PH003).....	112
Figure C3: Cross-method comparison of risk ratios (RR) for extreme (a) precipitation (19 stations) and (b) streamflow (28 gauges) events.....	113
Figure C4: Per-model spatial distributions of station-level risk ratios (RR) derived from the GCM- based attribution method for each of the four CMIP6 models.....	114
Figure C5: Per-model return-level curves and risk ratios (RR) for annual maximum 27-day mean flow at the Whitemouth River gauge (05PH003).....	115
Figure C6: As Figure C5 but for the Raven process-based hydrological model forced with each of the four CMIP6 models.....	116
Figure C7: Diagnostic basis for the regional attribution signal in 29-day annual maximum precipitation.....	117
Figure C8: Sub-regional Theil-Sen trends in physical-driver variables for the 1950 to 2022 period.	118

Chapter 1 Introduction

1.1 Background and Motivation

1.1.1 Flood Risk in Cold-Region Basins

Flooding remains one of the most damaging natural hazards globally, causing major loss of life and substantial economic damage over recent decades (Doocy et al., 2013; Kundzewicz et al., 2014; Tellman et al., 2021). In high-latitude regions, including much of Canada, flood generation is strongly shaped by cold-season snow accumulation and the spring transition, such that the timing and phase of water input can be as important as its total magnitude (Bush & Lemmen, 2019; Gray & Male, 1982; He et al., 2023). In these environments, the spring freshet concentrates the majority of annual runoff into a short period; snowmelt over frozen soils has historically accounted for over 80% of annual runoff in Canadian Prairie basins (Dumanski et al., 2015; Fang & Pomeroy, 2007). This concentration increases exposure to high flows, damaging communities, infrastructure systems, and agricultural landscapes (He et al., 2023; Stadnyk & Déry, 2021).

Cold-region spring flooding rarely results from a single meteorological trigger (Buttle et al., 2016; Gray & Male, 1982). Rather, it emerges from the interaction of snowpack magnitude, melt rate, rainfall on snow, soil freezing, and freeze-thaw cycles (He et al., 2023; Stadnyk & Déry, 2021). Frozen or partially frozen soils can greatly reduce infiltration, while low-relief floodplains may either attenuate or amplify runoff depending on their storage state and hydraulic connectivity (Bonsal et al., 2019; Hayashi et al., 2016; He et al., 2023). These interacting controls make both prediction and attribution more difficult than in predominantly rainfall-driven systems (Kundzewicz et al., 2014; National Academies of Sciences, 2016).

In Canada, major spring floods have repeatedly produced some of the country's most costly disasters, including the 2013 flood in southern Alberta and the 2021 flood in British Columbia, where prolonged precipitation over snow-affected and saturated catchments caused extensive damage (Bonsal et al., 2019; Bush & Lemmen, 2019; Canadian Climate Institute, 2024; Gillett et al., 2022; Stadnyk & Déry, 2021). Manitoba has likewise experienced several high-impact spring floods, including the major events of 1950, 1997, and 2011, which caused widespread inundation, evacuations, and multi-billion-dollar damages (Bonsal et al., 2019; Stadnyk & Déry, 2021). The 2022 flood in southern Manitoba falls within this broader history of, high-impact spring flooding, but it also raises a more specific scientific question: what caused the flood in a cold-region hydrologic system whose governing processes are highly nonlinear and increasingly exposed to anthropogenic climate change (He et al., 2023; Stadnyk & Déry, 2021).

1.1.2 Prairie Hydrology and Snow-Dominated Systems

The Canadian Prairies form a distinctive hydrologic setting characterized by low relief, widespread glacial depressions, and a strongly continental climate, all of which shape spring flood behavior by controlling how much snowmelt water is stored on the landscape and how quickly it reaches the channel network (Gray et al., 2001; Hayashi et al., 2016; He et al., 2023). Much of southern Manitoba consists of clay-rich, low-gradient glaciolacustrine plains where infiltration is limited and drainage is slow, especially during snowmelt (He et al., 2023; Stadnyk & Déry, 2021). In this landscape, snow redistribution by wind, accumulation in surface depressions, and the timing of melt affect both the total volume of water available for runoff and the spatial extent of hydrologically connected source areas (Fang & Pomeroy, 2007; Hayashi et al., 2016; He et al., 2023).

In southern Manitoba, the Prairie processes interact with additional regional controls, including transitions from agricultural plains to Precambrian Shield terrain, regulated rivers, and engineered flood-control infrastructure designed to protect Winnipeg and downstream communities (Bonsal et al., 2019; Stadnyk & Déry, 2021). The resulting flood response is highly heterogeneous across basins and strongly conditioned by relatively small differences in snow storage, melt timing, and spring precipitation (He et al., 2023; Stadnyk & Déry, 2021).

1.1.3 Anthropogenic Climate Change and Extreme Precipitation

Anthropogenic climate change is altering the frequency and intensity of extreme precipitation globally and across Canada, with direct implications for flood risk (Bush & Lemmen, 2019; Kundzewicz et al., 2014; Otto, 2023). Thermodynamic theory indicates that a warmer atmosphere can hold more water vapor, thereby increasing the potential for heavy precipitation, while circulation changes may further modify storm tracks and persistence (National Academies of Sciences, 2016; Otto, 2023). In this context, detection refers to identifying a statistically significant change, or signal, in an observed climate variable of any kind, while attribution refers to assigning that detected change to underlying causes such as anthropogenic forcing (National Academies of Sciences, 2016; Stott et al., 2016). Applied to extremes, such analyses have established human influence on temperature and precipitation extremes in many regions, including increases in heavy rainfall intensity (Kirchmeier-Young et al., 2025; Philip et al., 2020).

In Canada, recent studies reported warmer temperatures, changing precipitation regimes, and shifts in snow cover, especially in high-latitude and interior regions (Abdelmoaty et al., 2025; Bush & Lemmen, 2019; Vincent et al., 2015). Over the Prairies, projections indicate continued warming, possible increases in cold-season precipitation, and changes in both heavy rainfall and drought frequency, with implications for flood and water-scarcity risk alike (B. Bonsal et al., 2019;

He et al., 2023; Sinha et al., 2025; Stadnyk & Déry, 2021). These changes alter not only the magnitude of hydroclimatic forcing, but also its seasonal timing and phase, including the likelihood that rainfall will coincide with snowmelt and soil thaw (Hayashi et al., 2016; He et al., 2023; Stadnyk & Déry, 2021). For flood studies, the key challenge is not simply whether precipitation extremes are intensifying, but how changes in precipitation, snowpack accumulation and melt, soil moisture, and surface storage translate into streamflow response. This distinction is central to event attribution, particularly for flood events in cold-region basins.

1.1.4 Challenges in Flood Attribution

Event attribution studies aim to quantify how human influence has altered the probability or intensity of specific extreme events, including heatwaves, droughts, and heavy rainfall (National Academies of Sciences, 2016; Otto, 2023). Probabilistic event attribution frameworks typically compare the likelihood of an event in a factual climate, with observed anthropogenic forcing, to that in a counterfactual climate representing natural variability without human influence (Otto, 2023; Philip et al., 2020). Because the counterfactual climate cannot be observed directly, these frameworks rely either on climate-model simulations run with and without anthropogenic forcing, such as those coordinated under the Detection and Attribution Model Intercomparison Project (DAMIP; Gillett et al., 2016), or on observation-based statistical methods that use historical warming as a covariate. These approaches have advanced rapidly and are now guided by established protocols for probabilistic extreme-event attribution, such as the World Weather Attribution protocol (Philip et al., 2020), which emphasize objective event definition, observational analysis, climate-model evaluation, and synthesis across multiple attribution methods (Otto, 2023).

Flood events, particularly in cold-region basins, present specific challenges for probabilistic event attribution. First, many studies use precipitation as a proxy for flood hazard, because precipitation is simulated directly by climate models whereas discharge is not, and because hydrologic translation introduces an additional layer of uncertainty associated with hydrologic-model structure, parameters, and calibration (National Academies of Sciences, 2016; Otto, 2023; Winter et al., 2020). Second, for snowmelt- and rain-on-snow-driven floods, flood magnitude depends on antecedent snowpack, frozen-soil state, and basin wetness, so changes in precipitation extremes do not map proportionally onto changes in flood probability (Hayashi et al., 2016; He et al., 2023; Zscheischler & Seneviratne, 2017). Third, the compound nature of such events means that univariate attribution of any single driver can understate or overstate the anthropogenic influence on the flood itself (Zscheischler et al., 2018).

Attribution frameworks that incorporate streamflow therefore require a hydrologic translation step capable of representing basin-scale response to observed and modeled climate forcing. In this thesis, that role was filled by LSTM-based hydrologic models. Because several regional streamflow records also contain multi-year gaps that would otherwise weaken the extreme-value analysis underpinning attribution, the same LSTM framework was first used to reconstruct (gap-fill) incomplete records (Chapter 3) before being applied to climate-forced simulation (Chapter 4). Even so, important challenges remain, including LSTM generalization under nonstationary climate conditions, representation of rare extremes, and the propagation of uncertainty from climate forcing through streamflow simulation into the attribution metrics used to quantify anthropogenic influence (Kratzert et al., 2019; Martel et al., 2025; Wi & Steinschneider, 2022, 2024).

1.2 *Research Objectives*

The overall objective of this thesis is to quantify the influence of anthropogenic climate change on the likelihood of the 2022 spring flood in southern Manitoba by combining probabilistic event attribution frameworks for precipitation and streamflow. Observed precipitation, temperature, streamflow, and snowdepth records, together with CMIP6-simulations were used to evaluate how human influence has altered the likelihood of prolonged heavy spring precipitation and associated multi-week high flows, and how any changes in event likelihood relate to the compound snow, temperature, precipitation, and antecedent-wetness conditions that generated the flood. To achieve this overall objective, the thesis pursues four specific objectives (Appendix Table A2):

1. **Event characterization and rarity assessment:** Define the meteorological and hydrological 2022-type event using an objective, duration-optimized moving-window (RXnD) procedure, in which the accumulation window and its timing are selected to minimize the event exceedance probability rather than being fixed a priori, and quantify the historical rarity of the event using extreme value theory.
2. **Multi-method probabilistic attribution:** Apply and compare three probabilistic event attribution approaches, namely an observation-based scaling method with temperature as a covariate, an observational time-slice method based on change-point detection, and a GCM-based method where scenario without anthropogenic influence is considered, to estimate anthropogenic influence on both precipitation and streamflow extremes relevant to the event.
3. **Hydrological modeling for attribution:** Develop and evaluate LSTM-based hydrologic models for gap-filling incomplete streamflow records and simulating Whitemouth streamflow under factual and counterfactual conditions.

4. **Compound hydroclimatic drivers analysis:** Quantify and interpret the role of flood-generating conditions, specifically snowpack storage and spring temperature conditions, through joint-probability analysis.

1.3 Thesis Structure

This thesis comprises five chapters. Chapter 1: the introduction of the thesis, including the background on cold-region flood attribution, the research gap, and the overall and specific objectives. Chapter 2: describes the study region and datasets, and outlines the preprocessing, event-definition, and extreme-value framework used in the later chapters. Chapter 3: presents the LSTM-based streamflow gap-filling framework, including the comparison of Thiessen-based and ERA5-based meteorological forcing. Chapter 4: presents the manuscript, which applies multi-method probabilistic event attribution to the precipitation, streamflow, and compound drivers of the Spring 2022 Manitoba flood. This manuscript has been submitted to a peer-reviewed journal and is currently under review. Chapter 5: summarizes the thesis research and outlines the major conclusions and recommendations for future work. A unified reference list and four appendices (supplementary tables and figures, LSTM diagnostics, extreme-value and attribution diagnostics, and copula diagnostics) follow Chapter 5.

Chapter 2 Study Area, Data, and Preprocessing Framework

2.1 Study Area

The study domain (Figure 1a) covers southern Manitoba and extends slightly into adjacent northwestern Ontario, corresponding to the area covered by the hydro-meteorological network used in this thesis. This domain was selected because it encompasses the region most affected by the 2022 spring flood and contains the long-term meteorological and hydrometric networks required for event characterization and extreme-value analysis. The domain lies within the Canadian Prairie region and exhibits a hydroclimatic regime shaped by strong seasonality, low-relief terrain, and the dominant role of snow accumulation and melt in annual runoff generation. Much of the study region overlies the former basin of glacial Lake Agassiz, which left behind broad lacustrine plains with clay-rich soils and very gentle slopes (Last & Teller, 2004). These physiographic conditions favor ponding, slow drainage, and strong sensitivity to antecedent moisture and frozen-ground conditions during the spring melt season (Fang & Pomeroy, 2007; He et al., 2023). The study region drains toward Lake Winnipeg through several major river systems and numerous sub-basins (Government of Manitoba, 2021). Toward the east, the landscape transitions toward the Precambrian Shield, where steeper slopes, thinner soils, and more efficient drainage produce different runoff responses. Together, these contrasts produce a hydrologically diverse study region in which runoff response varies markedly among basins (Hayashi et al., 2016; He et al., 2023; Stadnyk & Déry, 2021).

The climate of the study region is strongly continental, with long cold winters, short transition seasons, and warm summers. In many basins, the annual hydrograph is dominated by the spring freshet, when the accumulated seasonal snowpack is released over a relatively short

period and generates a large fraction of annual runoff (Atashi et al., 2023; Boluwade & Rasmussen, 2015). The timing and magnitude of spring discharge depend on end-of-winter snow storage, melt rate, rainfall during the melt period, and the infiltration state of the soil surface. Because frozen or partially thawed soils restrict infiltration and increase runoff efficiency, similar water inputs can produce very different hydrologic responses depending on antecedent basin conditions (Fang & Pomeroy, 2007; Gray & Male, 1982; He et al., 2023). This snowmelt-dominated hydroclimatic regime explains why flood generation across the study region is closely tied to the spring transition. In some years, a deep snowpack followed by rapid warming is sufficient to generate major runoff even without exceptional rainfall. In other years, rainfall during the melt period becomes the dominant amplification mechanism, especially when it occurs on existing snowpack or over saturated soils (Musselman et al., 2018). Rain-on-snow conditions are particularly important because they combine direct liquid input with enhanced melt energy and often coincide with limited infiltration capacity. Flood behavior in this southern portion of Manitoba is therefore shaped not by meteorological forcing alone, but by the interaction of snowpack, melt timing, rainfall, and soil state (Buttle et al., 2016; He et al., 2023).

Southern portions of the domain are predominantly agricultural, with extensive artificial drainage and modified channels that influence runoff timing and routing (Manitoba Agriculture, 2016). In contrast, northern and eastern areas contain more forested landscapes with greater natural storage and evapotranspiration (Last & Teller, 2004). Spatial variability in drainage density, basin slope, soil properties, and land use produces heterogeneous hydrological responses across the study region. Low channel gradients and limited conveyance capacity in the clay plains can prolong flood peaks and extend floodplain inundation (Manitoba Infrastructure, 2023).

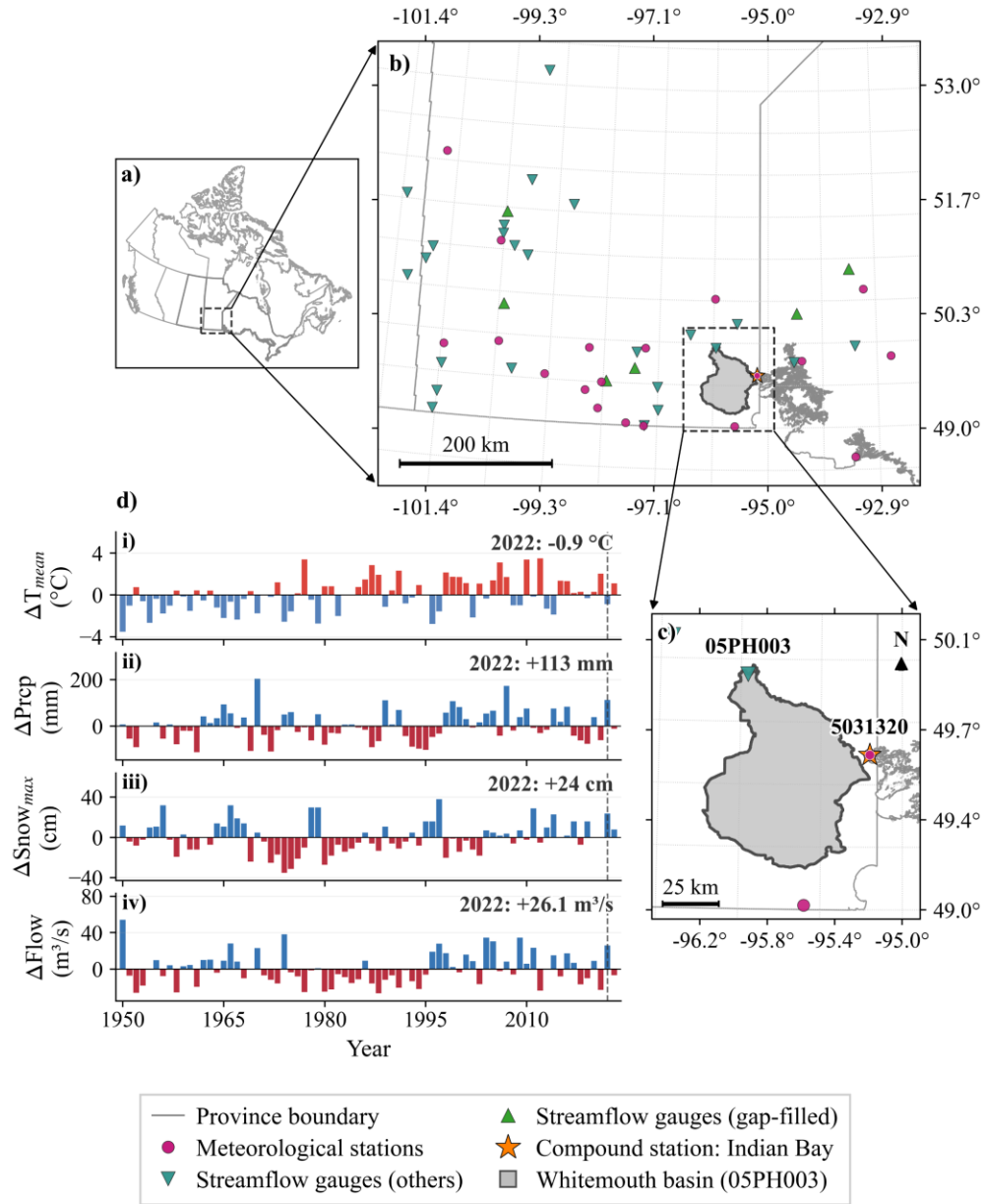


Figure 1: Study area and hydroclimatic context for the 2022 spring flood in southern Manitoba.

(a) Location of the study domain within Canada. (b) The study domain, showing the 19 meteorological stations, 28 streamflow gauges (six of them gap-filled), the Indian Bay compound-analysis station, and the Whitemouth River sub-basin (05PH003), the focal catchment for streamflow attribution (symbols as in the legend). (c) Detailed view of the Whitemouth sub-basin. (d) Seasonal anomalies of (i) March–June mean temperature, (ii) March–June precipitation, (iii) November–March peak snow depth, and (iv) March–June streamflow relative to the 1950–2023 mean, with 2022 marked by the dashed line.

These regional characteristics were especially evident during spring 2022. The flood developed within this regime, primed by winter snow accumulation and early-spring conditions. Rather than arising from a single short-duration storm, the event evolved through sustained hydroclimatic forcing during late April and May, when rainfall episodes overlapped with snowmelt and wet antecedent conditions. The 2022 event is therefore best understood as a compound spring flood, in which precipitation persistence, melt-season timing, and catchment wetness acted together, rather than as a purely rainfall-driven or purely snowmelt-driven event (Zscheischler et al., 2018). This context is important for interpreting the later attribution results because it clarifies why changes in precipitation extremes do not necessarily translate linearly into changes in streamflow extremes.

2.2 Data

This thesis draws on observed hydroclimatic records, reanalysis products, and climate-model simulations.

2.2.1 Streamflow Observations (HYDAT)

Observed daily streamflow data were sourced from the Water Survey of Canada HYDAT database, maintained by Environment and Climate Change Canada, which provides daily discharge records and station metadata for hydrometric stations across Canada. These observed discharge records serve as the training and evaluation targets for the LSTM models in Chapters 3 and 4. Within the study region, gauges were screened using the completeness criteria described in Section 2.3.1; 28 HYDAT streamflow gauges were retained (Table 1), of which six with substantial seasonal gaps were designated as the gap-filling targets of Chapter 3. The Whitemouth River gauge 05PH003 is the focal hydrometric station in this thesis. It was selected for the GCM-based

streamflow attribution for three reasons: it lies in the eastern portion of the domain that experienced the most severe 2022 flooding; its sub-basin contains the Indian Bay meteorological station, where the objective event definition of Chapter 4 identifies the most extreme 2022 precipitation accumulation; and it combines a long, nearly complete discharge record (0.5% missing days) with those co-located meteorological observations.

2.2.2 Precipitation and Temperature Observations

Meteorological station data were drawn primarily from Environment and Climate Change Canada (ECCC) and from the Serially Complete Dataset for North America (SCDNA; Tang et al., 2020). The ECCC network provides the long-term observed daily precipitation and temperature records in the region, while SCDNA provides serially complete daily precipitation, minimum temperature (Tmin), and maximum temperature (Tmax) records constructed from multiple observation sources and infilling procedures for 27,276 stations across North America from 1979 to 2018. A combination of SCDNA and ECCC precipitation and temperature data was used: SCDNA covers 1979 to 2018, and ECCC covers 1949 to 1979 and 2019 to 2023; the 1949 start supplies the 365-day antecedent forcing window required by the LSTM lookback (Chapter 3), while all analyses (Chapter 4) are reported for 1950 to 2023. This merged record is necessary in the study region because extreme-value analysis requires long, continuous series, whereas missing values and discontinuities complicate the direct use of any single raw source. A total of 19 precipitation and temperature stations were selected after quality control and completeness screening (Section 2.3.1). Among these, the Indian Bay station (5031320) was used for the compound snow-temperature analysis (Section 2.2.4).

Table 1: Summary of selected HYDAT streamflow gauges used in this study, including basin area, location, period of record, and percentage of missing days during March-October (1950-2023).

Station ID	River Name	Basin Area (km ²)	Latitude (°N)	Longitude (°W)	Period of Record	Missing Mar-Oct (%)	Regulation Status
05PH003	Whitemouth River Near Whitemouth	3,750	49.9388	95.9568	1950-04-13 to 2023-10-31	0.5	Regulated
05JM015	Cutarm Creek Near Spy Hill	766	50.5904	101.7429	1950-03-26 to 2023-10-31	1.2	Regulated
05KL001	Saskatchewan River At Grand Rapids	406,000	53.1639	99.3489	1950-09-12 to 2023-10-31	8.0	Regulated
05LH005	Waterhen River Near Waterhen	55,100	51.8485	99.5462	1951-03-01 to 2023-10-31	2.3	Regulated
05LJ005	Ochre River At Ochre River	348	51.0515	99.7881	1950-04-01 to 2023-10-31	1.2	Regulated
05LJ007	Turtle River Near Laurier	955	50.9488	99.5267	1950-03-30 to 2023-10-31	1.6	Regulated
05LJ010	Valley River Near Dauphin	2,880	51.2825	100.0165	1950-04-16 to 2023-10-31	1.8	Regulated
05LJ012	Vermilion River Near Dauphin	676	51.1875	100.0161	1950-04-01 to 2023-10-31	1.1	Regulated
05LJ025	Mossy River Below Outlet Of Dauphin Lake	8,740	51.4503	99.9711	1952-06-19 to 2023-10-31	12.5	Regulated
05LM001	Fairford River Near Fairford	79,800	51.5873	98.7118	1955-10-01 to 2023-10-31	8.3	Regulated
05MD004	Assiniboine River At Kamsack	13,000	51.5649	101.9166	1950-04-01 to 2023-10-31	1.2	Regulated
05MD005	Shell River Near Inglis	1,970	50.9633	101.3188	1950-04-23 to 2023-10-31	1.4	Regulated
05ME001	Assiniboine River Near Russell	19,400	50.8101	101.4354	1950-04-13 to 2023-10-31	3.7	Regulated
05MF001	Little Saskatchewan	2,610	50.3588	99.9075	1959-03-01 to 2023-10-31	12.2	Regulated

	n River Near Minnedosa						
05MJ001	Assiniboine River At Headingley	162,000	49.8683	97.4053	1950-03-01 to 2023-10-31	0.0	Regulated
05NF002	Antler River Near Melita	3,220	49.0573	101.0489	1950-03-01 to 2023-10-31	0.7	Regulated
05NG001	Souris River At Wawanesa	61,100	49.5969	99.6786	1950-03-22 to 2023-10-31	0.6	Regulated
05NG003	Pipestone Creek Near Pipestone	4,240	49.5960	100.9629	1950-03-01 to 2023-10-31	1.1	Regulated
05OC001	Red River At Emerson	102,000	49.0037	97.2238	1950-03-01 to 2023-10-31	0.0	Regulated
05OD001	Roseau River Near Dominion City	5,020	49.1912	96.9845	1950-03-01 to 2023-10-31	0.2	Regulated
05OE001	Rat River Near Otterburne	1,420	49.4619	97.0092	1950-03-01 to 2023-10-31	0.7	Regulated
05OF003	Boyne River Near Carman	1,130	49.5175	97.9446	1956-04-01 to 2023-10-31	20.4	Regulated
05OG001	La Salle River Near Sanford	1,800	49.6800	97.4321	1956-04-01 to 2023-10-31	21.3	Regulated
05PF063	Winnipeg River At Slave Falls	126,000	50.2250	95.5708	1950-03-01 to 2023-10-31	2.8	Regulated
05QC001	Chukuni River Near Ear Falls	4,920	50.8738	93.4848	1962-09-26 to 2023-10-31	18.1	Regulated
05QD006	Wabigoon River Near Quibell	6,490	49.9578	93.4005	1953-10-01 to 2023-10-31	5.5	Regulated
05QE009	Sturgeon River At Outlet Of Salvesen Lake	1,530	50.3522	94.4664	1960-06-16 to 2023-10-31	15.4	Regulated
05SA002	Brokenhead River Near Beausejour	1,580	50.0901	96.4283	1950-04-01 to 2023-10-31	1.2	Regulated

2.2.3 *ERA5 Reanalysis*

In addition to station data, this thesis uses the ERA5 reanalysis (Hersbach et al., 2020), the fifth-generation global atmospheric reanalysis produced by the European Centre for Medium-Range Weather Forecasts, which provides temporally continuous hourly fields on a 0.25° grid from 1940 to the present. In this thesis, ERA5 supplied the basin-scale meteorological forcing adopted for the LSTM gap-filling workflow of Chapter 3, following the forcing comparison presented there: fields of 2 m air temperature, total precipitation, and 10 m wind components were extracted from 1949 onward, so that the 365-day lookback window of the LSTM (Section 3.2) was fully populated from the first analysis year, 1950. ERA5 has been shown to perform comparably to station-based observations for hydrological modeling across much of North America, with the exception of the eastern United States, where observations lead to consistently better performance (Tarek et al., 2020).

2.2.4 *Snow Depth Observations*

Snow-related information was derived from a combination of the Canadian Historical Daily Snow Depth Dataset (CHDSDD), covering 1950 to 31 July 2017, and recent ECCC observational records starting 1 August 2017. Snow depth was used to characterize end-of-winter storage and to interpret the timing and progression of spring melt, both of which are key controls on runoff generation in Prairie basins (Gray & Male, 1982; Hayashi et al., 2016; He et al., 2023). Although the spatial coverage of snow observations is more limited than that of precipitation and temperature, available records indicate substantial variability in snow accumulation and melt across the Prairie-Shield transition, with important implications for runoff generation and flood risk. After the quality-control and completeness screening described in Section 2.3.1, the Indian

Bay station (5031320) was retained as the snow analysis station: its snow record is long and of high quality, it offers the co-located meteorological observations noted above, and it lies within the Whitemouth sub-basin, where the most extreme 2022 precipitation accumulation occurred (Chapter 4).

2.2.5 *CMIP6 Climate Model Simulations*

In the PEA framework for the GCM-based method, Coupled Model Intercomparison Project Phase 6 (CMIP6) simulations were considered (Eyring et al., 2016). Daily precipitation and temperature were extracted from a subset of CMIP6 models selected for their ability to reproduce the seasonal cycle and key climatological features of precipitation over the study region. The model set included CNRM-CM6-1, BCC-CSM2-MR, IPSL-CM6A-LR, and NorESM2-LM (Appendix Table A1), all available through the Earth System Grid Federation (World Climate Research Programme, 2025). Simulations from the 'Historical' experiment for 1950 to 2014, extended to 2022 with the SSP5-8.5 scenario, were treated as the factual climate. This experiment is driven by both natural and anthropogenic forcings: time-evolving greenhouse gas concentrations, anthropogenic aerosol emissions, stratospheric ozone changes, and land-use change, together with observed solar irradiance variability and volcanic stratospheric aerosols. Simulations from the 'HistNat' experiment of the Detection and Attribution Model Intercomparison Project (DAMIP) for 1950 to 2020 were used as the counterfactual climate. This experiment retains only the natural forcings (solar and volcanic forcings) from the historical simulations (Gillett et al., 2016). Because the two experiments are run with identical model physics and resolution, the difference between them isolates the effect of anthropogenic forcing on the simulated climate. The datasets used in this thesis are summarized in Table 2.

No bias correction was applied to the CMIP6 outputs, as the Probabilistic Event Attribution (PEA) framework focuses on changes in event probability rather than absolute magnitudes. Systematic biases arising from model physical parameterizations are consistent across both factual and counterfactual ensembles and thus cancel out in ratio-based attribution metrics such as the Risk Ratio (Kumari et al., 2019; Srivastava et al., 2020).

Table 2: Summary of meteorological, reanalysis, snow-related, and climate-model datasets used in this thesis.

Dataset	Variables	Time Coverage	Spatial Form	Primary Use
ECCC + SCDNA	PRCP, TMIN, TMAX	(1949-2023) ¹	Station-based (19 stations)	Event definition, extreme value analysis, observational attribution
HYDAT (WSC)	Daily discharge	1950-2023	Point-based (28 gauges; Whitemouth: 05PH003)	Event definition, hydrologic attribution, LSTM training and evaluation
ERA5 reanalysis	Temperature (2m), Precipitation, Wind (u10, v10)	1949-2023	Gridded (aggregated to basin-scale means)	LSTM gap-filling forcing (adopted after the Chapter 3 comparison)
CHDSDD + ECCC	Snow depth	1950-2023	Station-based (1 station: Indian Bay)	Snowpack characterization, compound driver analysis
CMIP6 (4 models)	Precipitation, Temperature	1950-2022 (Hist); 1950-2020 (HistNat)	Gridded (station-matched)	Probabilistic event attribution (factual vs. counterfactual climate), RR and FAR estimation

¹ 1949 supplies the 365-day LSTM lookahead (Section 3.2); analyses are reported for 1950–2023.

2.3 Data Preprocessing and Quality Control

Before conducting extreme-value analysis or attribution, the various observational, reanalysis, and model datasets were processed onto a common daily time base and, where required, aggregated to common spatial units (stations or basin means) to produce directly comparable time series.

2.3.1 Dataset Screening and Quality Control

HYDAT records were reviewed for record length, continuity, and regulation status, using the regulation flag reported in the HYDAT station metadata. All gauges retained after screening carry a regulated designation in that metadata; the screening therefore acted on completeness. Gauges with more than 20% missing daily values over the 1950–2023 analysis period were excluded from the core set, and among the 28 retained gauges, the six with more than 10% missing days in the hydrologically relevant March to October season were designated as the gap-filling targets of Chapter 3 (Table 1). Records failing these thresholds would undermine event characterization and LSTM-related analysis. Similarly, precipitation and temperature stations were screened for continuity and coverage. Internal gaps in the station-based climate series are reduced by the serially complete structure of SCDNA, which retains quality-controlled station observations and fills missing values by merging estimates from neighboring stations and from reanalysis products.

All relevant datasets were brought to a common daily temporal scale, since daily values form the basis of the event metrics used in the thesis. For ERA5, hourly data were aggregated to daily series, with precipitation accumulated to daily totals and atmospheric variables such as temperature and wind converted to daily means.

2.3.2 Spatial Processing

Two spatial processing steps were used, one for station-scale analysis and one for basin-scale analysis. For the station-based precipitation attribution, each of the 19 meteorological stations was paired with the climate-model grid cell whose center lies closest to that station, and the model series from that cell represented the station in the analysis. For basin-scale hydrologic forcing, gridded inputs were instead averaged over entire sub-basin polygons: ERA5 fields were clipped to the boundaries of the seven sub-basins used in Chapter 3 (the six gap-filling targets listed in Table 5 plus the Whitemouth sub-basin, 05PH003) and converted to area-weighted daily basin means. The purpose of basin averaging is to match the meteorological input to the hydrologic unit that generates the streamflow, rather than to a single point. This matters in the study region because precipitation and temperature vary spatially, so a single station can misrepresent the water input a basin actually receives.

2.4 Event Definition and Extreme Value Framework

This thesis adopts a probabilistic event-attribution perspective in which the 2022 flood is treated as a rare hydroclimatic event whose likelihood may differ between a factual climate and a counterfactual climate; that change in likelihood is quantified by the risk ratio and the fraction of attributable risk. The key concepts are introduced here because they shape how the data were prepared and how the event is interpreted, while the attribution metrics, uncertainty quantification, and method implementations are documented in Chapter 4; the complete workflow is summarized in Appendix Figure A2.

2.4.1 *Event Definition*

The event definition determines which extremes the attribution analysis compares. Because the 2022 flood was characterized by prolonged wet conditions and sustained high flow rather than a single-day peak, the meteorological component of the event was defined using multi-day precipitation accumulations and the hydrological component using multi-day mean discharge. Daily precipitation series were converted into moving-window accumulations over durations of 1 to 30 days (RXnD), and daily streamflow series into moving-window mean discharge over the same durations. These metrics preserve event persistence and allow precipitation and streamflow extremes to be analyzed in the same way. The optimization procedure used to select the most representative windows is presented in Chapter 4.

2.4.2 *Extreme Value Theory*

Once an event is defined, its rarity is estimated using extreme-value theory. In this thesis, annual maxima of the selected precipitation or streamflow were extracted, and a Generalized Extreme Value (GEV) distribution was fitted to those maxima (Coles, 2001). The return levels were calculated for each accumulation duration and corresponding return periods were obtained to characterize how unusual the 2022 event was in the observed record and then to compare how its probability changes between factual and counterfactual conditions.

Chapter 3 LSTM-Based Streamflow Gap-Filling Framework

3.1 *Motivation for Gap Filling*

In the study region, many streamflow records are sufficiently long to support climatological and hydrologic analysis, but they are not uniformly complete. Missing periods arise from delayed station commissioning, discontinuous operation, equipment failures, and changes in measurement programs. Such gaps are particularly problematic here because the most hydrologically informative period is concentrated in the spring freshet and early growing season, when snowmelt, rainfall, and antecedent wetness combine to produce the largest runoff events (Fang & Pomeroy, 2007; Hayashi et al., 2016). Even moderate missingness during this period can reduce the reliability of annual maxima and multi-day event metrics, and hence of inferred flood rarity. This concern is amplified here because the 2022 event was defined using multi-day mean discharge rather than single-day peaks (Section 2.4.1). The framework developed in this chapter therefore serves two roles in this thesis: it reconstructs incomplete hydrometric records for the regional extreme-value analysis, and it provides the trained-model basis for the CMIP6-driven streamflow attribution at Whitemouth in Chapter 4.

The challenge is not simply to interpolate short gaps. Several stations in the study region contain extended discontinuities that cannot be filled credibly using local smoothing or linear imputation. Gap filling was therefore treated here as a rainfall-runoff reconstruction problem in which missing streamflow is inferred from meteorological forcing and basin attributes. Long short-term memory networks are well suited to this task because they can learn nonlinear relationships between hydroclimatic inputs and daily streamflow while retaining memory of antecedent basin conditions through recurrent states (Gauch et al., 2021; Kratzert et al., 2019). In Prairie basins,

where streamflow response depends on cumulative wetness, seasonal snow conditions, and storage-connectivity effects, this capacity to represent temporal dependence is especially relevant (Hayashi et al., 2016; He et al., 2023). Because the reconstructions feed both the extreme-value analysis and the Chapter 4 attribution model, the choice of meteorological forcing matters throughout this chapter, motivating the comparison of station-based and reanalysis forcing that follows.

3.2 LSTM Modeling Framework

The gap-filling framework was implemented using the NeuralHydrology library (Kratzert et al., 2022) with the cudalstm model structure. A regional configuration was adopted for both experiments, so that each model could learn shared hydrologic behavior across its training basins while retaining basin-specific information through static attributes. This follows current large-sample hydrology practice, where regional LSTM models often outperform separate single-basin calibrations because the shared training set exposes the network to a wider range of hydroclimatic states and runoff responses (Gauch et al., 2021; Kratzert et al., 2019; Martel et al., 2025). The full architecture and training configurations of both experiments are summarized in Table 3, and the complete sets of dynamic inputs and static attributes are listed in Table 4. Model performance was evaluated using the Nash-Sutcliffe efficiency (NSE; Nash & Sutcliffe, 1970), the Kling-Gupta efficiency (KGE; Gupta et al., 2009), and the root mean square error (RMSE).

3.2.1 Thiessen-Based Forcing Configuration

The Thiessen-forcing experiment used dynamic meteorological inputs derived from Thiessen-polygon interpolation of in situ station observations to each sub-basin, together with a suite of static physiographic and climatic attributes describing basin-specific controls on

streamflow generation. In this approach, each sub-basin was partitioned into polygons that assigned every location the record of its nearest meteorological station, and the polygon area fractions provided the weights used to compute the basin-mean daily series (Thiessen, 1911). Hyperparameters were selected through an 81-member grid search over hidden size, batch size, output dropout, and sequence length, evaluated by the median validation NSE across basins (Appendix B, Figure B3); the selected configuration, with a hidden size of 256, a batch size of 64, an output dropout of 0.3, and a 365-day sequence length, achieved the highest median validation NSE of the grid.

Table 3: Architecture and training hyperparameters of the regional LSTM models under the Thiessen and ERA5 forcing configurations.

Hyperparameter	Thiessen configuration	ERA5 configuration
Model / library	cudalstm (NeuralHydrology v1.12.0)	
Hidden size	256	512
Sequence length (days)	365	
Output head	Regression, linear activation	
Dynamics embedding	FC 256-128-64, tanh, dropout 0.25	FC 256-128-64, tanh, dropout 0.25
Statics embedding	FC 32-64, tanh, dropout 0.2	
Output dropout	0.3	
Initial forget-gate bias	3	
Loss function	NSE	
Optimizer	Adam (Kingma & Ba, 2014)	
Learning-rate schedule (epoch: rate)	0: 2×10^{-4} ; 20: 1×10^{-4} ; 40: 5×10^{-5} ; 55: 2×10^{-5}	
Batch size	64	128
Epochs	60	
Training / validation / test periods	1950 to 2016 / 2017 to 2020 / 2021 to 2023	

Table 4: Dynamic meteorological inputs and static basin attributes used in the Thiessen and ERA5 LSTM configurations.

A check mark indicates the input is used in that configuration.

Input	Thiessen	ERA5
<i>Dynamic meteorological inputs</i>		
Daily total precipitation	✓	✓
Daily mean, maximum, and minimum 2 m air temperature	✓	✓
Daily potential evapotranspiration	✓	✓
Daily mean 10 m zonal wind (u), meridional wind (v), and wind speed		✓
<i>Static basin attributes</i>		
Basin polygon area (km ²)		✓
Drainage area at gauge (km ²)	✓	✓
Basin perimeter (km)	✓	
Compactness	✓	
Centroid longitude and latitude	✓	
Gauge longitude and latitude		✓
Mean annual precipitation, wet-day precipitation, and Dry-day fraction	✓	✓
Heavy-precipitation-day fraction		✓
Precipitation seasonality	✓	✓
Mean daily minimum and maximum temperature	✓	
Mean daily temperature		✓
Annual temperature range, Temperature seasonality	✓	✓
Snowfall fraction	✓	✓
Freezing-day fraction, Hot-day fraction		✓
Mean potential evapotranspiration		✓
Potential evapotranspiration seasonality, Aridity index (PET/P)		✓
Mean 10 m wind speed, Wind-speed seasonality		✓

3.2.2 *ERA5-Based Forcing Configuration*

The ERA5 forcing-comparison experiment followed the same general modeling philosophy but used a separate configuration tailored to the gridded forcing inputs (Table 3). Hyperparameters were selected through the same 81-member grid-search procedure (Appendix B, Figure B4). The search selected a larger hidden size of 512 for this configuration, consistent with the larger input set, which includes additional wind-related dynamic inputs and a broader suite of static attributes (Table 4), and the 365-day, batch-128 configuration was chosen from the top-ranked candidates to retain the sequence length used in the Thiessen experiment. Basin-scale meteorological forcing for this experiment was derived from ERA5 reanalysis (Section 2.2.3). Hourly fields of 2 m air temperature, total precipitation, and the 10 m zonal and meridional wind components were extracted, aggregated to daily time series, and converted to area-weighted basin means over the seven sub-basins considered in this chapter (the six gap-filling target sub-basins plus Whitemouth).

3.2.3 *Comparison Design*

The two forcing experiments differ in both data provenance and spatial representation. Thiessen-polygon forcing is directly grounded in observations and reflects conventional station-based basin forcing. ERA5 forcing is spatially continuous and predictor-rich but may inherit biases associated with reanalysis dynamics and observation assimilation, especially in data-sparse regions (Hersbach et al., 2020; Tarek et al., 2020). The aim here is not to identify one universally superior dataset, but to evaluate how these differences affect streamflow reconstruction skill and to determine which forcing is most appropriate for the reconstruction task. The Thiessen-based model was trained regionally on all 28 gauged sub-basins, following large-sample practice in which additional training basins improve regional LSTM skill (Kratzert et al., 2019), whereas the

ERA5-based model was trained on the seven sub-basins for which gridded basin forcing was prepared (the six gap-filling targets plus the Whitemouth sub-basin, 05PH003). Skill was evaluated for both models on these common seven sub-basins over identical validation and test periods.

3.3 Gap-Filling Implementation

Six stations exceeded the 10% March to October missingness threshold (Section 2.3.1) and were identified as the gap-filling targets: 05LJ025, 05MF001, 05OF003, 05OG001, 05QC001, and 05QE009. As summarized in Table 5, these stations combined high March to October missingness (12.16% to 21.26%) with multi-year discontinuities of 3 to 12 consecutive years.

Table 5: Summary of the six-target gap-filled stations, including Station ID, March to October NaN percentage, fully missing years, longest consecutive gap, and missing-year runs.

Station ID	Station Name	March-October NaN (%)	Fully Missing Years	Longest Consecutive Gap (years)	Missing-Year Runs
05LJ025	Mossy River Below Outlet of Dauphin Lake	12.53	5	3	1950-1951; 1953-1955
05MF001	Little Saskatchewan River Near Minnedosa	12.16	9	9	1950-1958
05OF003	Boyne River Near Carman	20.41	10	6	1950-1955; 1997-2000
05OG001	La Salle River Near Sanford	21.26	13	7	1950-1955; 1959-1965
05QC001	Chukuni River Near Ear Falls	18.12	12	12	1950-1961
05QE009	Sturgeon River At Outlet of Salvesen Lake	15.44	10	10	1950-1959

The screening used the March to October season because that is the flood-relevant period, but the reconstruction itself was performed on full-year daily series. Reconstructing the full year

avoids introducing discontinuities at season boundaries and gives the LSTM the continuous input history its memory states require (Lees et al., 2022). Model skill was assessed on the withheld validation (2017–2020) and test (2021–2023) periods rather than by masking observed segments and predicting them, an approach consistent with other deep-learning reconstruction studies (Gauch et al., 2021; Wi & Steinschneider, 2022). The trained models were then applied to reconstruct the missing segments of each target record. Figure 2 and Figure 3 show the reconstructed streamflow over the missing periods of each station under the Thiessen and ERA5 forcing, respectively, and Figure 4 shows the epoch-wise validation performance of both models together with the epoch selected for each station.

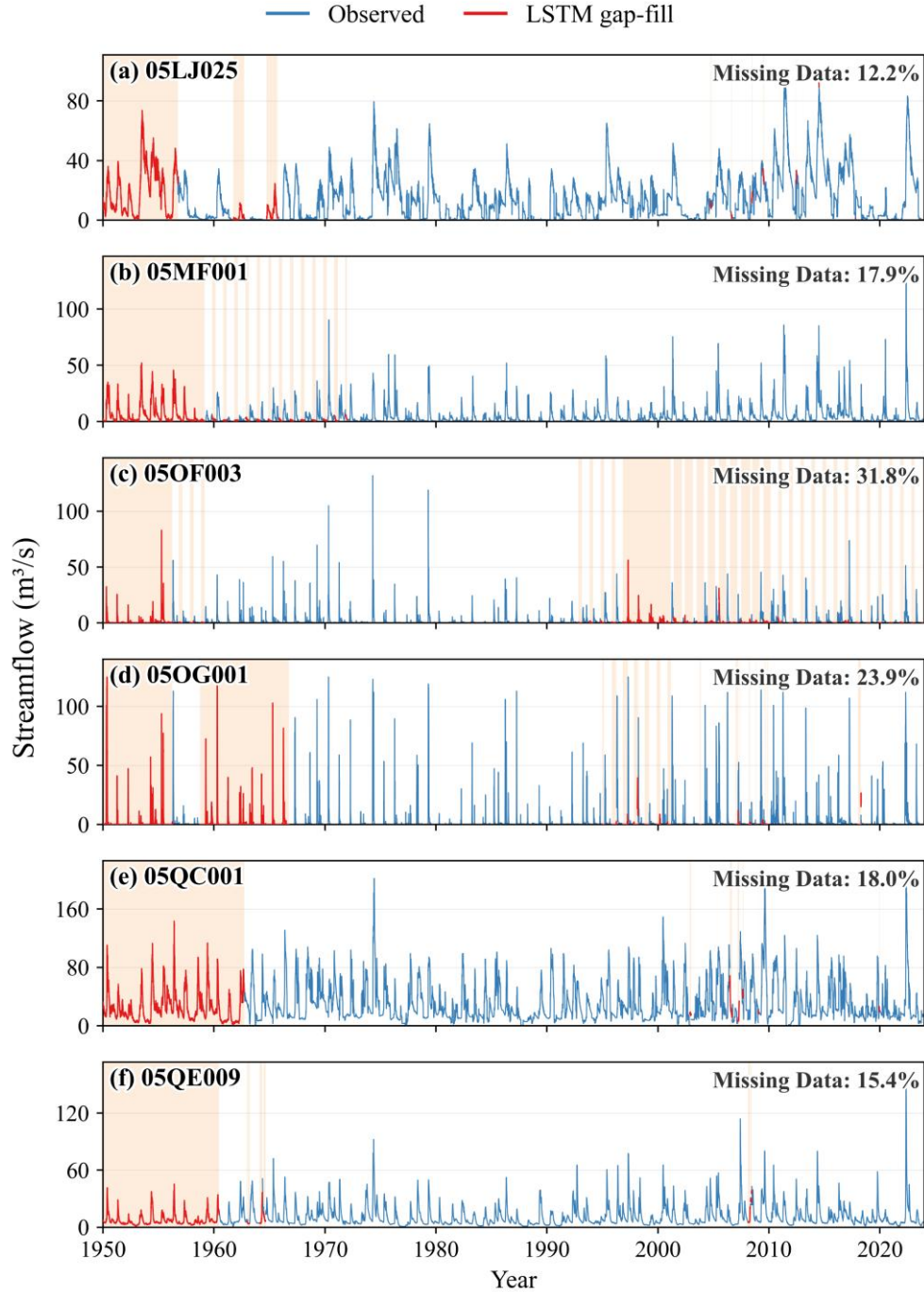


Figure 2: Gap-filling diagnostic for the six target stations under the Thiessen-forcing regional LSTM. Blue lines show the observed daily record; red lines show LSTM-simulated streamflow drawn only where observations are missing (shaded spans), with each panel annotated by its full-record missing fraction. Simulated values are taken from the training, validation, or test simulation covering each date, using the best validation-NSE epoch for each station.

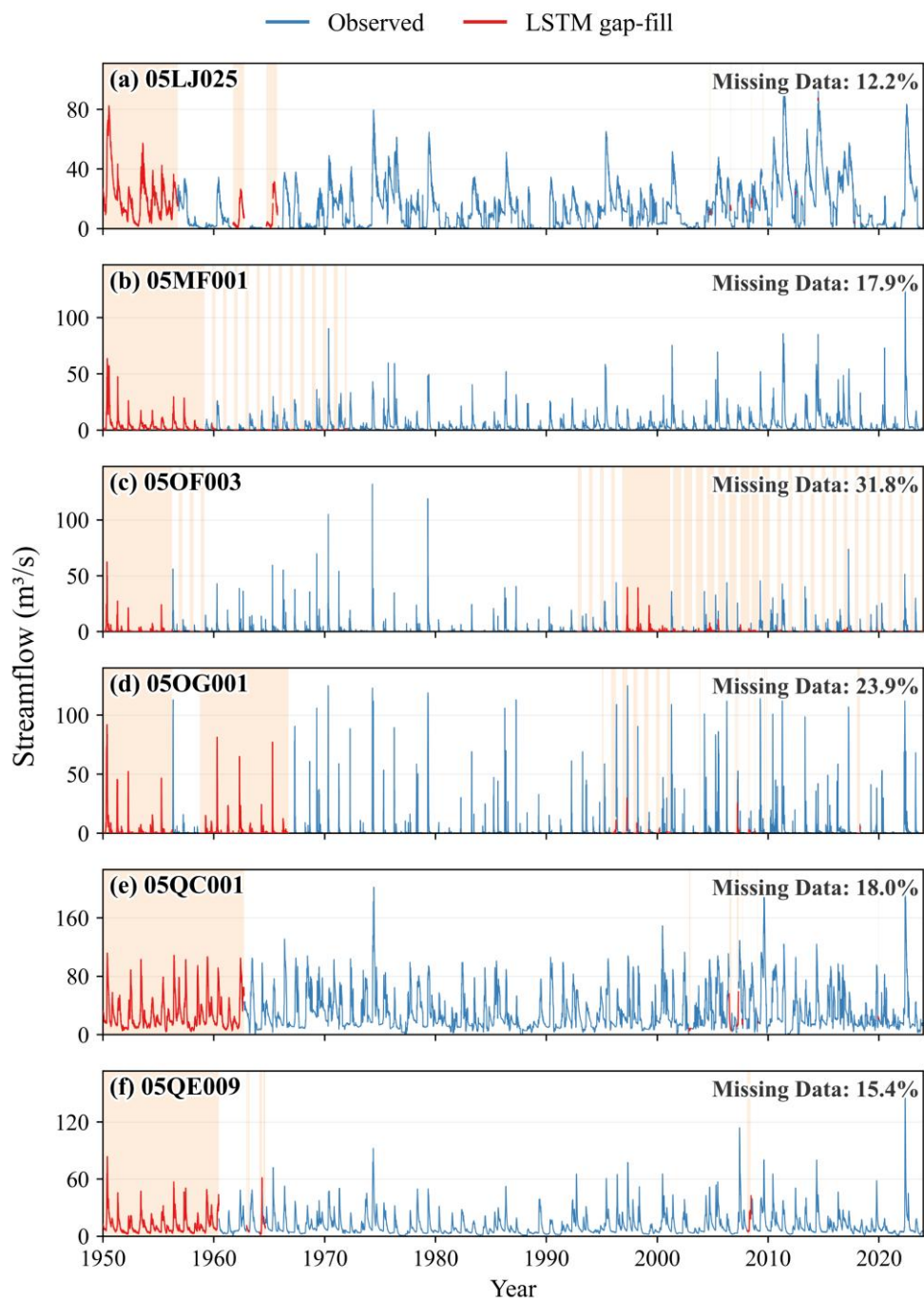


Figure 3: As Figure 2, but under the ERA5-forced regional LSTM adopted for the final gap-filled product.

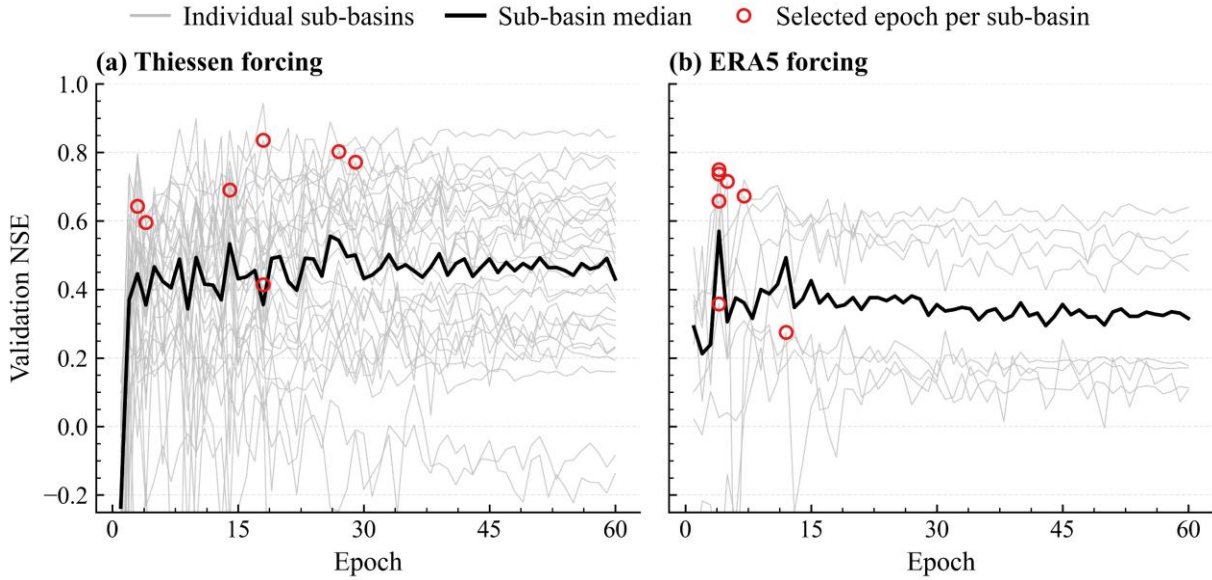


Figure 4: Epoch-wise validation performance of the two regional gap-filling LSTMs: (a) Thiessen and (b) ERA5 forcing. Thin gray lines show validation NSE for individual basins, the black line the basin median, and red circles the epoch selected for each focal station (maximum validation NSE, RMSE tiebreak), as used for the reconstructions in Figure 2 and Figure 3.

3.4 Results: Gap-Filling Performance

Across the regional experiments, both forcing strategies produced useful streamflow reconstructions, but their comparative performance shows a clear advantage for ERA5 reanalysis on the seven-sub-basin comparison set (Figure 5). The ERA5-driven model reached higher test-period NSE at most stations, and its largest gains occurred at the stations where the two forcings disagreed most. Because ERA5 was evaluated on the seven-sub-basin subset rather than on all 28 stations, this comparison is a forcing comparison for the gap-filling basins, not a replacement of the full Thiessen-based regional workflow.

The reconstructed series themselves (Figure 2 and Figure 3) behave physically plausibly over the missing periods: the infilled segments reproduce the strong seasonal structure of Prairie

streamflow, with spring freshet peaks followed by summer recession, and the magnitudes of the reconstructed annual peaks lie within the range of the observed record at each station rather than drifting toward the mean. The two forcings produce broadly similar reconstructions during low-flow periods and differ mainly in the magnitude of individual spring peaks, anticipating the skill differences quantified below. Figure 4 shows how validation skill evolved during training and which epoch was used for each station's reconstruction. Validation NSE rises rapidly over the early epochs before flattening, and the epoch of maximum validation skill differs across stations, which motivated the per-station epoch selection: using each station's best validation epoch, rather than a single global stopping point, provides a form of station-level early stopping that guards against overfitting the small validation sample. The selected epochs are generally earlier for the ERA5 model than for the Thiessen model, consistent with its larger batch size processing more data per epoch.

3.4.1 Station-Wise Performance Comparison

ERA5 achieved a higher test-period NSE at four stations: 05LJ025 (0.942 vs. 0.776), 05MF001 (0.835 vs. 0.606), 05PH003 (0.877 vs. 0.831), and 05QE009 (0.852 vs. 0.627). The NSE margins at 05MF001, 05QE009, and 05LJ025 are particularly large, with ERA5 exceeding Thiessen by 0.229, 0.225, and 0.166 NSE units, respectively; differences of this size represent a shift from moderate to strong reconstruction skill. Thiessen achieved a higher test-period NSE at the remaining three stations: 05OF003 (0.827 vs. 0.659), 05OG001 (0.843 vs. 0.809), and 05QC001 (0.773 vs. 0.757). These Thiessen margins are smaller, the largest being 0.168 NSE units at 05OF003 and the other two below 0.04. Where ERA5 performed better it generally did so by large NSE margins, whereas the Thiessen advantages were mostly minor.

3.4.2 *Forcing Selection*

The ERA5 forcing was retained for the final reconstructions on the strength of its test-period performance. The withheld 2021-2023 test period provides the cleanest basis for comparison: it remained unseen throughout model development, and performance on unseen data is the criterion most relevant to gap filling, since the missing segments themselves contain no observations. On this split, ERA5 achieved the higher skill on all three metrics. The mean NSE, which measures how well the simulated series reproduces the observed variability relative to the observed mean, was 0.819 for ERA5 against 0.755 for Thiessen. The mean KGE, which jointly evaluates correlation, bias, and variability, showed the largest relative advantage, at 0.759 against 0.643, with ERA5 higher at six of the seven stations. The mean RMSE, which expresses the typical magnitude of the simulation error in flow units, was correspondingly lower for ERA5, at 7.92 against 9.67 m³ s⁻¹. The agreement of all three metrics, each emphasizing a different aspect of simulation quality, indicates that the ERA5 advantage is a general property of the reconstruction rather than an artifact of any single evaluation measure. The structural properties noted in Section 3.2.3 also favor ERA5 for basin-scale forcing: representing meteorological conditions over the entire basin area rather than at a small number of point locations, so the area-weighted means capture spatial gradients in precipitation and temperature that a sparse station network can miss. Its complete coverage supplies consistent inputs across all sub-basins regardless of how many stations lie in or near each basin, and its wind components add predictors relevant to cold-region runoff generation. The distributional check in Section 3.5 shows no detectable change in the annual-maximum statistics used in the subsequent analysis.

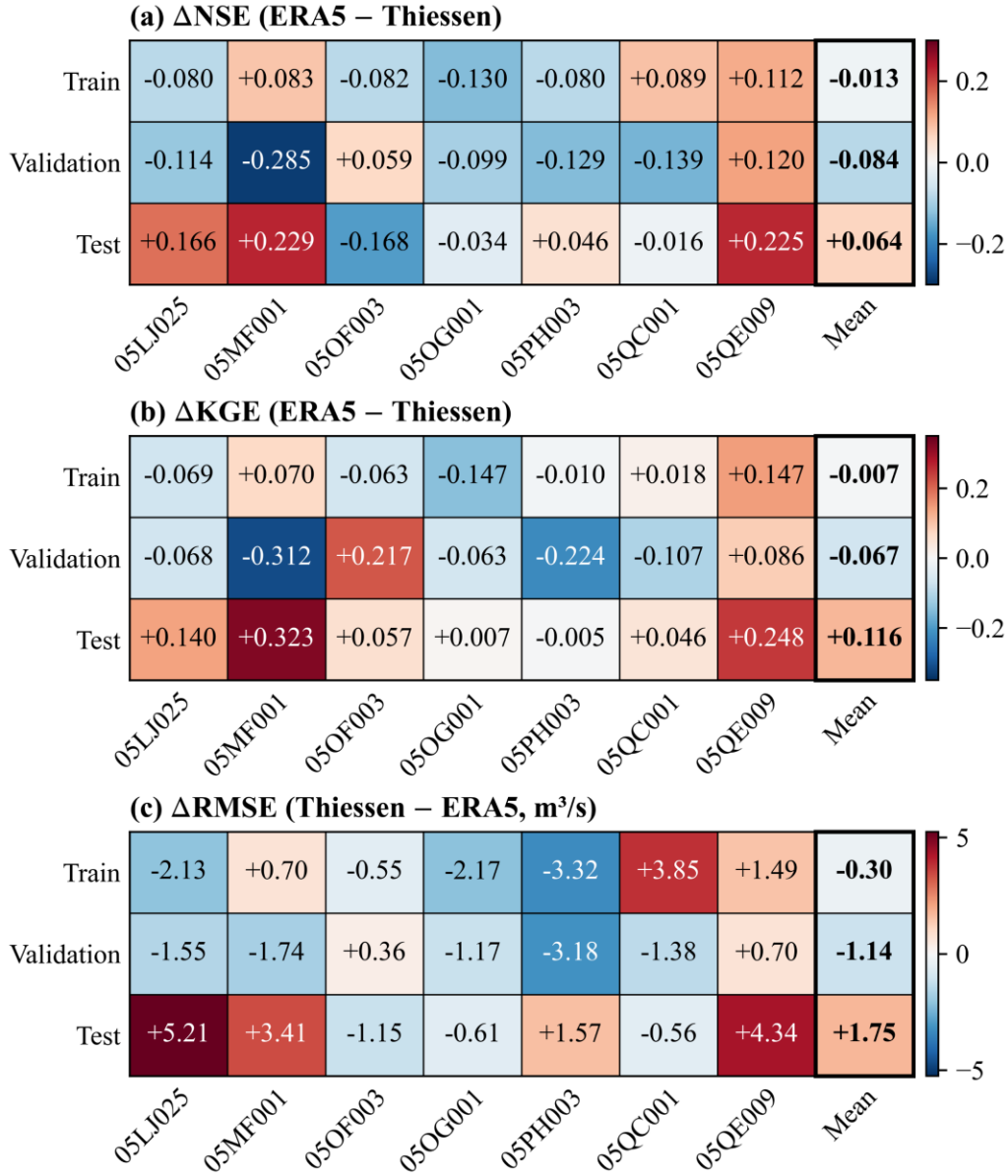


Figure 5: Station-wise comparison of ERA5 and Thiessen meteorological forcings for the regional LSTM across the seven study sub-basins, showing the difference in (a) Nash-Sutcliffe efficiency (Δ NSE, ERA5 – Thiessen), (b) Kling-Gupta efficiency (Δ KGE, ERA5 – Thiessen), and (c) root mean square error (Δ RMSE, Thiessen – ERA5, m³ s⁻¹) for the training, validation, and test periods. Signs are oriented so that positive values (red) indicate better ERA5 performance and negative values (blue) better Thiessen performance in every panel; the Mean column gives the average across the seven sub-basins.

3.5 *Implications for Subsequent Analyses*

The forcing comparison also carries implications for the CMIP6-driven streamflow generation objective at station 05PH003 (Chapter 4). Although ERA5 delivered the stronger overall reconstruction skill and was therefore retained as the forcing basis for the regional gap-filling workflow (Section 3.4.2), the single-basin Whitemouth attribution model faces a different set of constraints. Because only one catchment was modeled in the attribution step, no multi-station interpolation or spatial averaging across gauges was required, and the observed meteorological record at the co-located Indian Bay station (5031320) provided the most direct representation of the basin-mean forcing. The Chapter 4 attribution LSTM was therefore trained on this observed station record rather than on ERA5.

Beyond the forcing decision, the main question for this thesis is not whether every reconstructed daily value is exact, but whether gap filling preserves the statistics of flood-relevant streamflow extremes. This was checked directly by comparing the empirical cumulative distribution functions of annual maximum RXnD streamflow computed from the observed records and from the ERA5-forced gap-filled records adopted in Section 3.4.2, at all six target stations (Figure 6). The two sets of distributions are statistically indistinguishable, with two-sample Kolmogorov-Smirnov tests failing to reject equality at every station: across all 24 station-duration combinations, the KS statistic does not exceed 0.08 and the smallest p-value is 0.95, far from any conventional rejection threshold, while the annual-maximum sample at each station grows from 61-69 years to the full 74 years. Gap filling therefore did not materially alter the annual maxima or the RXnD metrics used later in the thesis. This is consistent with the broader finding that LSTM-based hydrologic models can reproduce important aspects of extreme-event behavior when trained on sufficiently informative input-output data (Gauch et al., 2021). The corresponding check for the

Thiessen-forced reconstructions yields the same conclusion (Appendix B, Figure B5), showing that the preservation of annual-maximum statistics does not depend on the choice of forcing.

The practical benefit of gap filling is a larger sample of annual maxima, not a change in their magnitudes. Before reconstruction, the six target stations were missing 12.2% to 21.3% of their March to October daily discharge records, including gaps of 3 to 12 consecutive years (Table 5). A year with missing discharge records during the spring months cannot provide an annual maximum, because Prairie annual maxima occur almost exclusively during the spring freshet; such years therefore drop out of the sample entirely. As a result, the stations contributed only 61 to 69 usable years out of the 74-year analysis period (1950 to 2023). After reconstruction, every station contributes all 74 years. A longer, uninterrupted sample reduces the standard errors of the fitted GEV parameters (Coles, 2001) and ensures that the estimated tail behavior reflects hydrologic variability rather than which years happened to be observed. For the multi-week spring extremes analyzed in this thesis, this gain of 5 to 13 years of annual maxima per station is the main contribution of the gap-filling work.

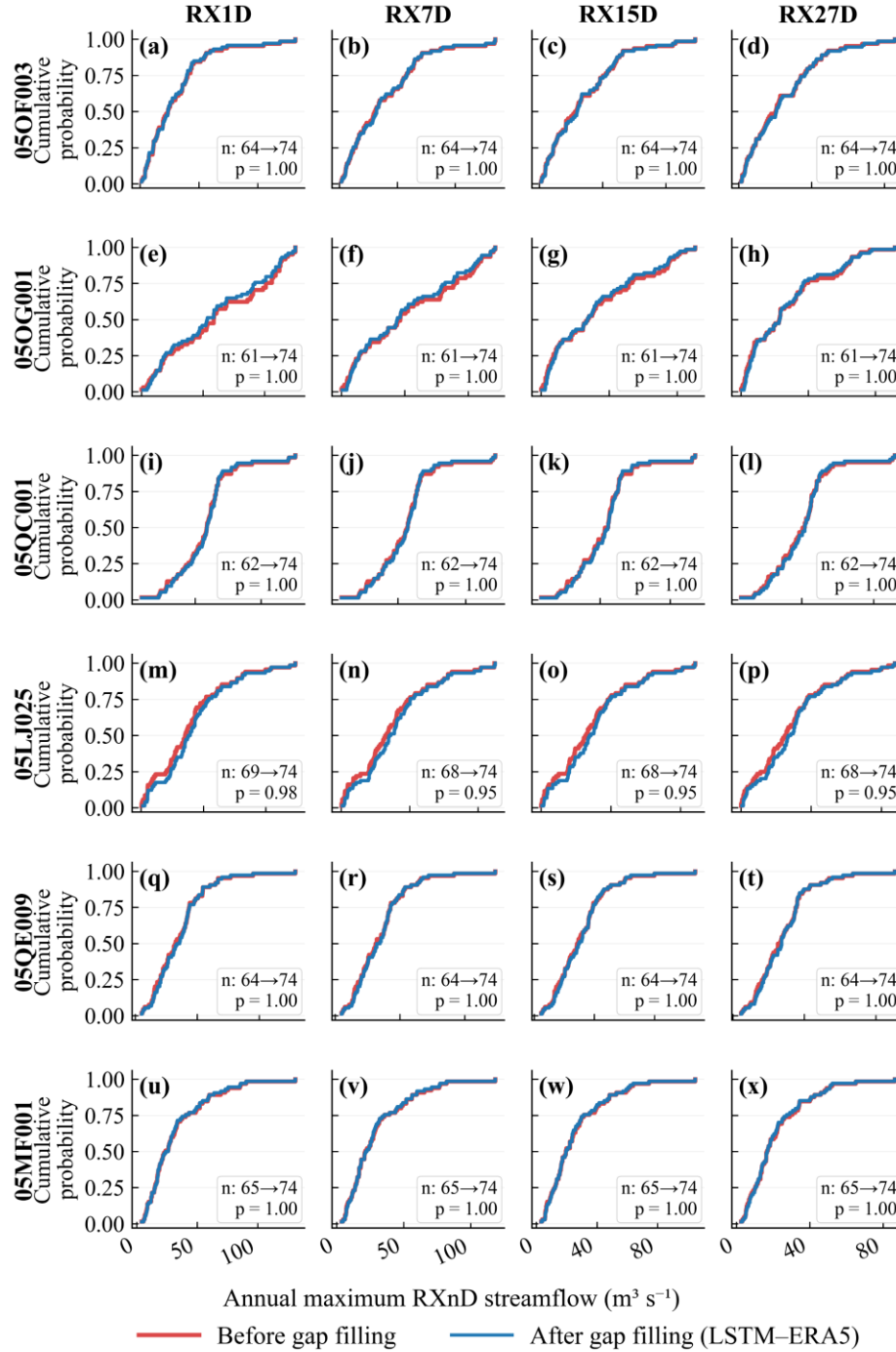


Figure 6: Empirical cumulative distribution functions (ECDFs) of annual maximum RXnD streamflow before (red) and after (blue) LSTM-based gap filling for the six gap-filling target stations. Columns correspond to accumulation durations of 1, 7, 15, and 27 days; rows correspond to individual HYDAT stations. Each panel reports the sample size before and after gap filling (n) and the two-sample Kolmogorov-Smirnov test statistic and p-value. All series are restricted to the March to October season over the 1950 to 2023 period.

Chapter 4 Assessing the Role of Climate Change in the 2022

Manitoba Spring Flood

Muhammad Zain Ul Abideen¹, Setegn Abate Getu¹, Chandra Rupa Rajulapati¹, Alex Crawford²

¹Department of Civil Engineering, University of Manitoba, Winnipeg, MB, CA; ²Department of Environment and Geography, University of Manitoba, Winnipeg, MB, CA.

Abstract

In spring 2022, southern Manitoba experienced a major flood, causing severe damage to infrastructure, agricultural land, and rural communities. Although spring flooding is recurrent in this region, the degree to which anthropogenic climate change is altering the likelihood and severity of such events remains poorly understood. We quantified the influence of anthropogenic factors on the Spring 2022 flood using a probabilistic event attribution framework. Risk ratios (RR) were calculated between factual and counterfactual climates using three methods: a scaling method perturbing observations by temperature between the 1901 and 2022 climates, a time-slice method splitting the record at a detected change point, and a GCM method based on historical and historical-natural CMIP6 scenarios. We further applied a bivariate copula framework to evaluate the joint influence of antecedent snow accumulation and April mean temperature. RR estimates from all methods indicated that precipitation magnitude was likely exacerbated by climate change, particularly in southeastern Manitoba, the portion underlain by the Canadian Shield rather than the lacustrine Red River Valley to the west. The signal for flood magnitude, however, was not statistically significant, highlighting the uncertainty that remains in attributing hydrometeorological events, including compound occurrences, in settings of high natural variability. These results suggest that warming may be increasing the background likelihood of

flood-favorable conditions, while flood magnitude in cold-region basins remains strongly modulated by antecedent snowpack and melt processes. This partial decoupling between precipitation and flood attribution has direct implications for cold-region adaptation planning.

Keywords: probabilistic event attribution; compound events; flood attribution; cold-region hydrology; Canadian Prairies; copula.

4.1 Introduction

The broader context for this chapter is established in Chapter 1 (Sections 1.1 and 1.2), which also defines the research gap that motivates this work; the probabilistic event attribution (PEA) framework is summarized in Chapter 2 (Section 2.4). The aim here is narrower: to apply that framework to the Spring 2022 event itself. Flood attribution in cold-region Prairie basins is particularly demanding because extreme runoff is generated through interacting meteorological and hydrological controls rather than precipitation alone, and because snow storage, melt timing, frozen-ground conditions, antecedent soil moisture, and depressional storage jointly modulate the translation of precipitation into discharge (Buttle et al., 2016; He et al., 2023; Merz et al., 2021; Musselman et al., 2018). Most existing PEA studies for floods use precipitation as a proxy for flood hazard, in part because climate models more readily simulate precipitation and precipitation-based event definitions are simpler to implement (Philip et al., 2020; van Oldenborgh et al., 2021). In snowmelt- and rain-on-snow-dominated basins, that framing may misrepresent the actual hazard when runoff is controlled by antecedent basin state rather than by rainfall alone (He et al., 2023; Zscheischler et al., 2018).

The Spring 2022 flood in southern Manitoba provides a demanding test case for this distinction. The event developed during a period of substantial antecedent snow storage, delayed melt under cooler early-spring conditions, and prolonged rainfall, producing multi-week rather

than single-storm hydroclimatic extremes. In southeastern Manitoba, the Whitemouth River sub-basin experienced sustained high discharge during late April and May, while nearby precipitation stations recorded unusually persistent accumulations. Local states of emergency were declared by 45 municipalities and nine First Nations, more than 2,000 residents of Peguis First Nation were displaced following ice-jam flooding along the Fisher River, and provincial infrastructure losses alone were estimated at approximately C\$193.5 million (Environment and Climate Change Canada, 2023; Manitoba Transportation and Infrastructure, 2023). The wider history of high-impact Manitoba spring floods (1950, 1997, and 2011) is reviewed in Chapter 1 (Section 1.1.1).

This chapter quantifies the role of anthropogenic climate change in the Spring 2022 extreme precipitation and flooding. Specifically, we (i) define the meteorological and hydrological characteristics of the event objectively and assess its historical rarity; (ii) evaluate how anthropogenic forcing has altered the probability of extreme precipitation and streamflow during spring using three attribution approaches (GCM-based, scaling-based, and time-slice-based methods); (iii) compare attribution outcomes across methods to assess inter-method consistency and characterize uncertainty; and (iv) examine the extent to which changes in precipitation extremes translate into flood magnitude in a cold-region basin influenced by snowmelt, rain-on-snow processes, and antecedent soil moisture. Streamflow under factual and counterfactual climates is generated using both a data-driven (LSTM) and a process-based (Raven) hydrological model at the Whitemouth sub-basin, building on the forcing evaluation and model development of LSTM presented in Chapter 3. A bivariate copula-based framework then evaluates whether the joint probability of flood-conducive antecedent snow storage and cool-spring melt-season conditions has changed between early and late observational periods. The overall workflow is summarized in Appendix Figure A2.

4.2 *Study Area and Data*

The study domain covers southern Manitoba and adjacent northwestern Ontario, with the Whitemouth River sub-basin (gauge 05PH003) as the focal catchment for streamflow attribution. The regional setting, the snowmelt- and storage-controlled flood regime, and the seasonal hydroclimatic context of the 2022 event are described in full in Chapter 2 (Section 2.1) and are summarized for reference in Figure 1. Only the elements directly required for the attribution analysis are restated here.

The observational and model datasets are documented in Chapter 2 (Sections 2.2 and 2.3). The attribution analysis used the 19 precipitation and temperature stations (ECCC and SCDNA), the 28 HYDAT streamflow gauges (including six gap-filled stations), and the Indian Bay snow depth record retained after the Chapter 2 screening, together with daily CMIP6 Historical and Historical-Natural precipitation and temperature. The six gap-filled streamflow series were reconstructed using the regional LSTM gap-filling framework developed in Chapter 3, forced with basin-averaged ERA5 reanalysis inputs. ERA5 forcing was adopted for the gap-filling step (Chapter 3, Section 3.4). Two further points specific to this chapter are noted. First, the GCM-based precipitation attribution used four CMIP6 DAMIP models (BCC-CSM2-MR, CNRM-CM6-1, IPSL-CM6A-LR, and NorESM2-LM; Appendix Table A1), selected for their skill in reproducing the precipitation climatology and seasonal cycle of the study region (Appendix Figure A1). Second, the streamflow attribution at Whitemouth used an LSTM trained on the observed daily record for that single sub-basin, with observed meteorological forcing from the co-located Indian Bay station (5031320); the model configuration and the rationale for observed rather than interpolated forcing are given in Section 4.3.4. All analyses focused on the April to May window, the season in which the 2022 event developed (Section 4.3.1).

4.3 *Methods*

We applied a probabilistic event attribution (PEA) approach based on three methods (scaling, time-slice, and GCM-based) to quantify anthropogenic effects on precipitation and streamflow (19 precipitation stations and 28 streamflow gauges, respectively; Figure 1), and a bivariate (compound) approach to assess the joint snow-temperature preconditioning at Indian Bay. The attribution metrics are the risk ratio (RR) and the fraction of attributable risk (FAR). Denoting by p_1 the probability of an event at least as extreme as the observed 2022 magnitude under the factual climate and by p_0 the corresponding probability under the counterfactual climate, $RR = p_1/p_0$ and $FAR = 1 - p_0/p_1$. Values of RR greater than one indicate that the event is more likely under present-day conditions, and positive FAR values indicate that part of the event probability is attributable to anthropogenic influence. The general PEA logic and extreme-value basis are introduced in Chapter 2 (Section 2.4); this section documents the event-definition procedure and the method-specific implementation details.

4.3.1 *Event Definition and Rarity*

Defining the event is critical in PEA analysis, as attribution results can be sensitive to the choice of spatial and temporal event definition (F. Otto et al., 2018; van Oldenborgh et al., 2021). Here we applied the objective method of Cattiaux & Ribes (2018), which provides an automated framework for determining the optimal spatial and temporal window for the event. For each station, all n -day (n from 1 to 30) moving windows within the April to May 2022 season were analyzed to extract RXnD metrics, defined as n -day accumulated precipitation and n -day mean streamflow. For each duration, an annual-maximum RXnD series was constructed from the historical record, and the exceedance probability of the 2022 event was estimated both empirically

and by fitting a stationary Generalized Extreme Value (GEV) distribution using maximum likelihood estimation (Coles, 2001), with goodness-of-fit evaluated using the Kolmogorov-Smirnov (Teegavarapu, 2019) and Anderson-Darling statistics (Stephens, 1974). The duration minimizing the GEV-based exceedance probability (that is, yielding the largest return period) was adopted as the station's representative extreme event duration. This approach reduces subjective bias and ensures spatial and temporal consistency across analyses (Leach et al., 2020; Ren et al., 2020).

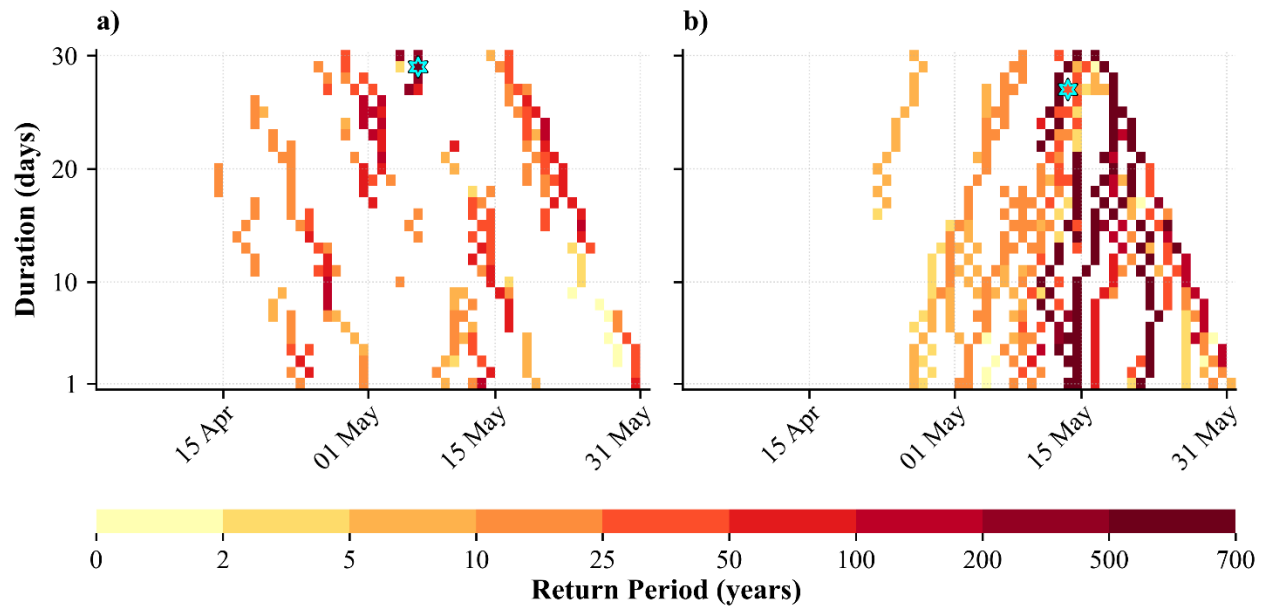


Figure 7: Event-based heatmaps of return periods estimated using the generalized extreme value (GEV) distribution for April to May 2022: (a) precipitation events and (b) streamflow events. Each cell represents an event defined by its duration (y-axis) and center date (x-axis), with color indicating the GEV-estimated return period in years. The star in (a) indicates the event with the highest return period among all precipitation stations; the star in (b) marks the corresponding streamflow station's highest return period event.

A 29-day accumulation from 22 April to 20 May 2022 at the Indian Bay station (49.62°N , 95.2°W) emerged as the most extreme precipitation window in the record. Over this period, 189.4 mm of rainfall was recorded, corresponding to the highest 29-day total in the roughly 73-year observational record. This translates to an empirical exceedance probability of about 1.37%

(approximately a 73-year return period), based on the station's ranked historical 29-day totals. A fitted GEV distribution placed this event even deeper into the tail, corresponding to roughly a 675-year return period (though with large uncertainty). We therefore defined this 29-day, station-centric window as the representative meteorological event for attribution. A heatmap of April to May 2022 precipitation rarity (Figure 7a) further illustrates that late April to mid-May contained the most anomalous multi-day rainfall clusters in the study region. The star marker in Figure 7a denotes the event with the highest return period among all precipitation stations, identifying Indian Bay as the location of the most extreme precipitation signal in the basin. The spatial and temporal concentration of high return periods confirms that the spring 2022 precipitation extreme was driven not by a single intense storm but by a persistent, multi-week wet spell extending from late April into late May.

Applying the same methodology to streamflow identified a corresponding hydrological extreme in spring 2022 within the Whitemouth River sub-basin. The most extreme window was a 27-day period from 30 April to 26 May 2022 at the Whitemouth River gauge (05PH003; 49.94°N, 95.96°W), during which the mean discharge reached $141.6 \text{ m}^3 \text{ s}^{-1}$. This prolonged high-flow episode corresponds to an exceedance probability of about 3.9% (approximately a 1-in-25.7-year event) based on the 1950 to 2023 record, with the GEV fit yielding a consistent estimate. We therefore adopted this RX27D window as the representative flood event. The timing of the hydrological extreme closely followed the precipitation anomaly. The streamflow heatmap (Figure 7b) shows that the highest return periods cluster in the late April to May interval, with peak-flow windows occurring shortly after the precipitation hotspot identified in Figure 7a. The star marker in Figure 7b denotes the return period of the corresponding streamflow event at the Whitemouth gauge rather than the basin-wide maximum. Although shorter-duration flow spikes occurred in

mid-May, they were less persistent; the selected multi-week window captured the dominant sustained flood signal of the season. GEV diagnostics confirm that the annual 27-day flow maxima at Whitemouth are well described by a GEV model (Appendix Figure C2), with the 2022 value among the highest on record but within the model’s expected range.

Considering the duration window, the extreme-value time series was constructed, and the risk ratio was obtained using the three methods detailed below. Uncertainty in the GEV parameters and all derived quantities was quantified by nonparametric bootstrap resampling (Efron, 1979; Wehner et al., 2024). Each extreme-value sample was resampled 5,000 times and the full attribution calculation repeated for every resample; where return levels or joint exceedance probabilities depend on fitted distributions, 10,000 Monte Carlo samples were drawn from each fit within every bootstrap resample, so that parameter uncertainty is carried through to the RR distribution (Appendix Figures C1, C2, and C8). This matters for multi-week spring extremes, where samples are small and rare-event probabilities are weakly constrained, conditions that can produce wide or asymmetric confidence intervals when the counterfactual probability is very small (Philip et al., 2020; Wehner et al., 2024).

4.3.2 Attribution Methods

4.3.2.1 Scaling Method

This approach estimates anthropogenic influence directly from the observed record by statistically scaling extreme values with long-term temperature changes. We fitted a nonstationary GEV to the annual maxima series (RX_nD) at each station, allowing the GEV location and scale parameters to vary as exponential functions of the smoothed global mean land temperature anomaly, following Clausius-Clapeyron-type scaling relationships (Philip et al., 2020), while the shape parameter was held constant. This formulation ensures a physically interpretable,

multiplicative response of extremes to warming. We smoothed the temperature time series with a 4-year running mean to filter out high-frequency variability before fitting (Dhasmana et al., 2023). From the fitted nonstationary GEV, we inferred two climate states: a 1901-like counterfactual climate representing early-twentieth-century conditions, and a 2022-like climate representing present-day conditions. We then derived the exceedance probabilities of the observed 2022 event magnitude under each climate state and computed the risk ratio as the ratio of the 2022 probability to the 1901 probability. This observations-only method provides an independent attribution check that isolates the thermodynamic response of extremes to warming and has been applied in recent regional precipitation- and flood-attribution studies (Dhasmana et al., 2023; Philip et al., 2020; van Oldenborgh et al., 2021).

4.3.2.2 *Time-Slice Method*

This approach compares the frequency of extremes between two time slices of the observed record, treating the earlier period as a proxy for the natural climate and the recent period as the current climate. We split the 1950 to 2023 record into two periods using an objectively identified change point detected with Pettitt’s test (Pettitt, 1979; Xie et al., 2013) applied to the time series of annual maximum RXnD, after smoothing the series with a 10-year moving window to emphasize low-frequency shifts (Dhasmana et al., 2023). The Pettitt test was performed separately for each variable and station, yielding breakpoints that clustered around the mid-1980s for both variables (median 1986 across stations; for example, 1982 at Winnipeg for precipitation and 1993 for Whitemouth streamflow). Where the Pettitt test did not identify a statistically significant change point, the record was split at its median year as a fallback. We designated the pre-change segment (for example, 1950 to 1981 for precipitation) as the baseline (counterfactual) climate and the post-change segment (for example, 1982 to 2023) as the recent (factual) climate. Stationary

GEV distributions were fitted independently to the annual maxima in each period, the exceedance probabilities of the observed 2022 event were evaluated against both distributions, and RR was computed analogously to the scaling method.

4.3.2.3 *GCM-Based Method*

We used climate-model simulations to estimate event probabilities under factual and counterfactual conditions, drawing on CMIP6 outputs for the Historical (Hist, natural and anthropogenic forcings, extended to 2022) and Historical-Natural (HistNat, natural forcings only) scenarios. For precipitation, we extracted the annual maxima of station-matched 29-day precipitation (RX29D) from 1950 to 2022 for each ensemble and fitted separate stationary GEV distributions to the Hist and HistNat maxima. The risk ratio is the ratio of the factual to the counterfactual exceedance probability of the observed 2022 magnitude. A confidence interval that excludes unity indicates a statistically detectable change in event probability at the stated confidence level (Philip et al., 2020; van Oldenborgh et al., 2021).

For streamflow, the analysis was carried out only at the Whitemouth sub-basin, which experienced the most extreme 2022 streamflow and precipitation conditions in the study region. We set up two hydrological models (a data-driven LSTM model and a process-based Raven model) at the Whitemouth River gauge 05PH003 (Section 4.3.4) and obtained annual maxima of 27-day mean discharge (RX27D) from the LSTM- and Raven-generated Hist and HistNat streamflow series.

4.3.3 *Bivariate (Compound) Attribution Framework (Snow-Temperature)*

We further assessed whether the joint occurrence of antecedent snow accumulation and cool spring conditions (one of the major reasons for the 2022 flooding) had changed between the early and late observational periods. This compound perspective is motivated by the physical

understanding that spring floods in cold-region prairie basins are governed not by a single driver but by the co-occurrence of snow accumulated over the winter and a melt driven by early-spring temperatures that synchronize runoff release with the onset of warm weather and rain-on-snow events (Buttle et al., 2016; Merz et al., 2021; Musselman et al., 2018). Two variables were selected to represent compound flood preconditioning. The first is antecedent snow accumulation, defined as the seasonal maximum daily snow depth observed during November to March (in cm), which represents the peak available snow water storage at the end of the winter accumulation season and is the primary quantity governing spring melt-driven runoff in this region. Maximum daily snow depth was preferred over cumulative snowfall because it characterizes the actual snowpack present on the ground after sublimation, mid-winter melt, and densification losses. The second is spring melt-season temperature, defined as the April mean of daily mean temperature, where daily mean temperature at each station was computed as the mean of daily maximum and minimum temperatures. Cooler April conditions delay melt onset and increase the probability that residual snow intersects subsequent warming and spring precipitation events, a mechanism consistent with the hydroclimatological drivers diagnosed for the Spring 2022 event (Section 4.5.2) and with rain-on-snow flood generation in continental cold-region basins more broadly (Musselman et al., 2018).

Marginal distributions of the snow and temperature variables were fitted independently. Three candidate distributions were evaluated for each variable: the normal, Student's *t*, and GEV distributions for temperature, and the Weibull, gamma, and lognormal distributions for peak snow depth. Model selection used a composite rank-sum criterion aggregating the Akaike Information Criterion (Akaike, 1974), the Kolmogorov-Smirnov statistic, and the mean squared error between the empirical and fitted cumulative distribution functions; the distribution achieving the lowest aggregate rank was selected as the best-fitting marginal for each variable, station, and period. The

bivariate structure between snow and temperature was modeled using copulas, with three candidate families evaluated at each station and period: the Gaussian, Frank, and Student's *t* copulas, selected to span elliptical, Archimedean symmetric, and mild upper-tail-concordant structures (Bevacqua et al., 2017; Salvadori et al., 2007; Zscheischler & Seneviratne, 2017). The best-fitting copula was selected using the same rank-sum criterion in uniform (pseudo-observation) space. Selected marginal distributions and copula families are listed in Appendix Table D1.

The observational record was divided into an early period (≤ 1985 ; $n = 36$ years) and a late period (> 1985 ; $n = 38$ years), consistent with the time-slice method applied in the univariate precipitation analysis. Because the physical condition of interest is cool-spring plus high-snow preconditioning, the joint exceedance probability was defined on the lower tail of temperature and the upper tail of snow, with the 2022 observed values used as thresholds. Operationally, temperature pseudo-observations were transformed so that cold extremes mapped to high values, allowing the joint exceedance probability to be evaluated directly from the fitted copula. A compound risk ratio was then computed as the ratio of the late-period joint exceedance probability to the early-period probability. Marginal risk ratios for temperature and snow were also computed separately to understand the relative contributions of each variable; the temperature marginal used the lower-tail definition, and the snow marginal used the upper-tail definition so that both tails represented flood-conducive extremes. Uncertainty was quantified with the bootstrap procedure of Section 4.3.1, applied to paired early- and late-period probabilities, with 90% confidence intervals. A compound $RR > 1$ indicates that the joint occurrence of 2022-like high-snow and cool-spring conditions became more likely in the recent period. The analysis was conducted at Indian Bay.

4.3.4 Hydrological Modeling

4.3.4.1 LSTM Streamflow Model at Whitemouth

For the Whitemouth River, an LSTM was configured within the same modeling framework as Chapter 3 (NeuralHydrology, cudalstm architecture), adapted to a single-basin setting in which static basin attributes were omitted since only one catchment was modeled. Because a single sub-basin involves no interpolation across multiple gauges, the observed record of the co-located station 5031320 was used directly as the basin-mean forcing rather than a spatially interpolated or reanalysis-based product, providing the most faithful representation of the catchment forcing for this gauge. The model was trained on observed daily streamflow from 1 January 1950 to 31 December 2010, validated over 1 January 2011 to 31 December 2016, and tested over 1 January 2017 to 31 December 2023. After training, model parameters were fixed and the LSTM was forced directly with CMIP6 Hist and HistNat precipitation and temperature series projected onto the Whitemouth sub-basin. This construction ensured that the basin-mean forcing during the CMIP6-driven simulations was dimensionally identical to that used during training, that runoff-generation relationships remained fixed across scenarios, and that Hist minus HistNat differences reflected the signature of anthropogenic forcing on the meteorological drivers rather than structural inconsistencies between the calibration and application forcing (Gauch et al., 2021; Kratzert et al., 2019). The observed and simulated daily streamflow for this single-basin configuration, shown alongside the Raven simulation, is presented in Figure 14.

4.3.4.2 Raven Streamflow Model at Whitemouth

In parallel, Hist and HistNat streamflow simulations were generated using the Raven hydrological modeling framework, a flexible, process-based system designed to represent watershed-scale hydrological processes through modular model components (Craig et al., 2020).

The Raven configuration was adopted from the HydroHub xSSA-NA framework, with basin physiography, land use and land cover, and soil representation retained from the baseline setup. Land use and land cover data were derived from the North American Land Change Monitoring System (NALCMS) 2010 land-cover product at 250 m resolution (Commission for Environmental Cooperation, 2010), and basin physiographic properties were obtained from HydroSHEDS. Daily basin-mean precipitation, minimum temperature, and maximum temperature were obtained from the station 5031320 record used for the LSTM configuration, ensuring a consistent observational forcing basis across both modeling approaches; observed daily streamflow at gauge 05PH003 was used as the calibration target. All inputs were converted into Raven-compatible time-series files. To minimize sensitivity to initial conditions, a warm-up period from 1 January 1950 to 31 December 1954 was applied and excluded from evaluation. Calibration was conducted over 1 January 1955 to 31 December 2010, followed by validation over 1 January 2011 to 31 December 2023. Parameter optimization was performed using the OSTRICH calibration toolkit (Tolson & Shoemaker, 2007) while preserving the conceptual structure of the adopted HydroHub configuration. After calibration and validation, the Raven model was forced with CMIP6 Hist and HistNat inputs from the four models, with calibrated parameters held fixed across all simulations so that differences in simulated streamflow reflected climate forcing rather than parameter changes.

4.3.4.3 *Model Evaluation and Diagnostics*

Performance of both the LSTM and Raven streamflow simulations was evaluated using standard hydrological skill metrics, including the Nash-Sutcliffe efficiency (NSE), Kling-Gupta efficiency (KGE), and root-mean-square error (RMSE), computed separately for the calibration, validation, and test periods appropriate to each framework. Diagnostic plots include epoch-wise

validation performance and flow-duration curves, provided in Appendix Figures B1 and B2, while observed and simulated hydrographs for both models are shown in Figure 14.

4.4 Results

4.4.1 *Scaling-Based PEA Analysis*

The scaling-based probabilistic event attribution (PEA) analysis indicates a mixed but predominantly anthropogenic signal in the Spring 2022 extreme across the 19 precipitation stations and 28 streamflow gauges examined in this study (Figure 8). At the point-estimate level, 13 of the 19 precipitation stations (about 68%) and 20 of the 28 streamflow gauges (about 71%) exhibit RR values at or above unity (Figure 8, center column), indicating that under a 2022-like climate the observed event magnitude is at least as likely as under the 1901-like counterfactual. At the 90% confidence level, 7 of the 19 precipitation stations and 5 of the 28 streamflow gauges produce confidence intervals lying entirely above unity, representing statistically detectable anthropogenic influence at these sites. Only one precipitation station (49.56°N, 99.07°W) yields a confidence interval entirely below unity, indicative of dominant natural variability, and no streamflow gauge produces a confidence interval entirely below unity. The remaining stations and gauges straddle $RR = 1$, reflecting the limited constraint on attribution available at the point scale from the observational record. The overall proportion of statistically significant detections is notably lower for streamflow (about 18%) than for precipitation (about 37%), an outcome consistent with prior attribution work over Canada showing weaker detectable forcing signals in hydrological extremes than in the meteorological extremes that drive them (Teufel et al., 2017).

A coherent east-west gradient is evident in the RR maps for both variables (Figure 8). In the central-south portion of the study area (approximately 49.0 to 50.0°N, 98.3 to 97.2°W), spanning the Red River Valley and the immediate surroundings of Winnipeg, point RR estimates

cluster in the natural-variability regime: the median point RR across the eight precipitation stations in this sub-region is 0.85; only three stations exhibit $RR \geq 1$; none are significant at the 90% level; and the adjacent streamflow gauges cluster near unity with limited statistical significance. In the eastern sub-region (approximately 95.5° to 92.7°W), covering the Winnipeg River basin and the Lake of the Woods region, the Canadian Shield portion of the study area characterized by near-surface bedrock, thin soils, and conifer-dominated forest cover, 6 of 7 precipitation stations exhibit point estimates at or above unity and 5 are significant at the 90% level; the majority of the detectable streamflow anthropogenic signals are also most pronounced in this eastern sub-region (Figure 8b). A similarly oriented but more modest anthropogenic signal is noted in western Manitoba (approximately 101.0 to 99.9°W), where 2 of the 3 available precipitation stations yield significant anthropogenic influence at the 90% level and adjacent streamflow gauges are consistent with this pattern. The stronger anthropogenic signal in eastern Manitoba, together with the natural-variability signature in the central-south sub-region, is broadly consistent with continental-scale analyses that report detectable anthropogenic contributions to Canadian precipitation over the long observational record, with more robust signals at larger spatial scales and higher latitudes and greater uncertainty at smaller regional scales (Kirchmeier-Young et al., 2025). This sub-regional contrast is reflected in the fitted GEV scaling coefficients themselves: the sub-regional median scaling coefficient across the eight central-south stations is -1.10 , with non-positive values at six of eight stations, in contrast to a median of 0.00 at the eastern stations and $+0.61$ at the western stations (Appendix Figure C7).

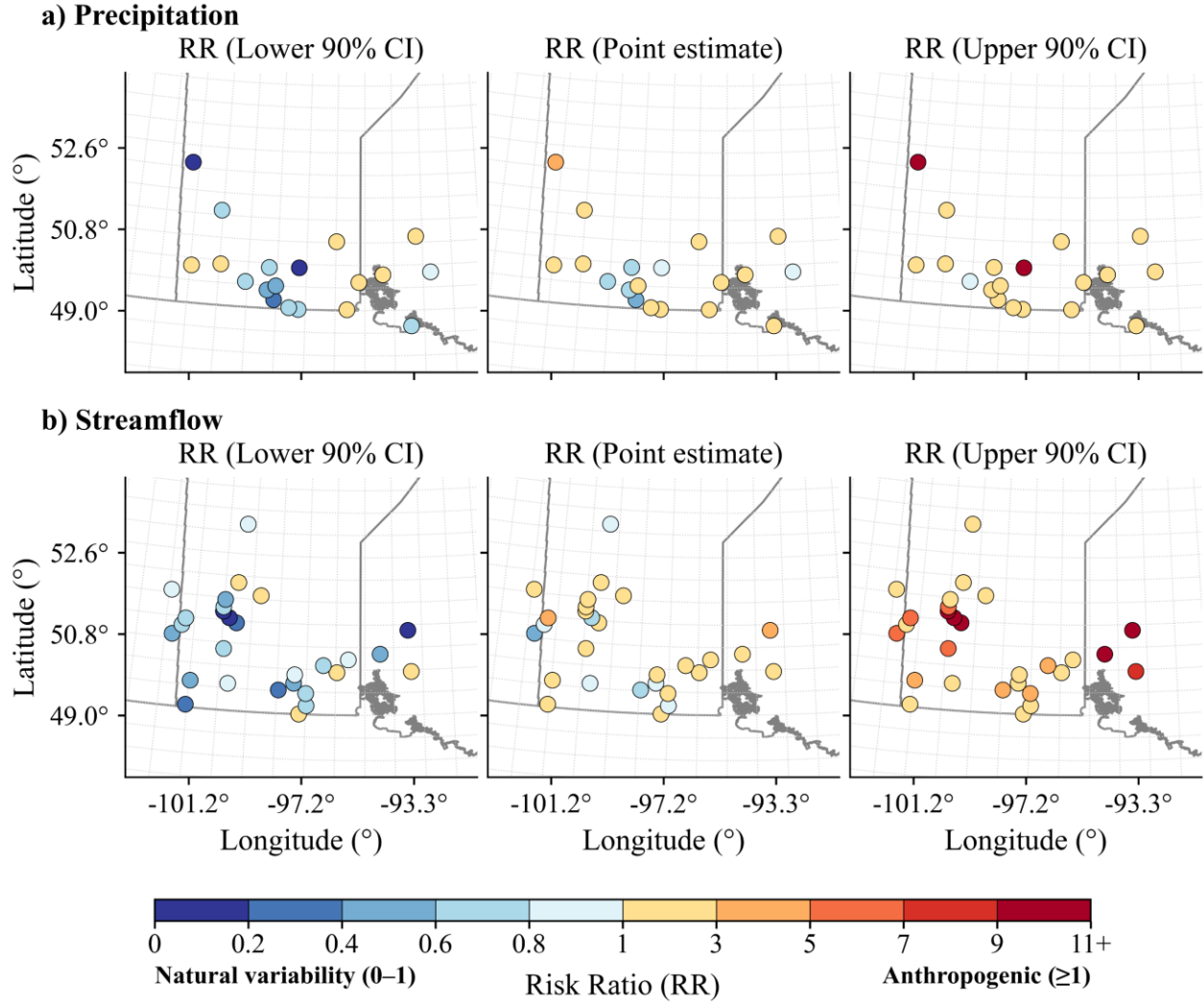


Figure 8: Spatial distribution of station-level risk ratios (RR) derived from the scaling-based attribution method. Panels show results for (a) precipitation (19 stations) and (b) streamflow (28 gauges). For each variable, the left, center, and right columns display the lower bound of the 90% confidence interval, the point estimate, and the upper bound of the 90% confidence interval, respectively. Colors denote RR magnitude, with blue shades ($RR < 1$) indicating dominance of natural variability and warm shades ($RR \geq 1$) indicating a central estimate of anthropogenic influence.

To illustrate the station-scale behavior underlying these spatial patterns, return-level curves were examined for a representative precipitation station and a representative streamflow gauge, both located within the anthropogenically influenced eastern sub-region (Figure 9). At the precipitation station at 49.02°N , 95.60°W , the 23-day accumulation of 162.3 mm observed in 2022

yields a point-estimate RR of 1.49 with a 90% confidence interval of [1.33, 1.69] (Figure 9a). Because the lower bound of the confidence interval exceeds unity, the increase in event probability is statistically significant at the 90% level: the observed magnitude is approximately 49% more likely under 2022-like temperatures than under the 1901-like counterfactual, corresponding to a fraction of attributable risk of about 0.33. The upward shift of the fitted GEV return-level curve between the factual and counterfactual climates is evident across the full range of return periods. At the Whitemouth River streamflow gauge (49.94°N, 95.96°W), which drains into the lower Winnipeg River, the scaling analysis yields a point-estimate RR of 1.00 with a 90% confidence interval of [1.00, 2.19] (Figure 9b). The lower bound anchored at unity indicates that no detectable reduction in the probability of the observed flood magnitude is supported by the bootstrap resampling, while the upper bound admits the possibility of an approximate doubling of risk. The attenuated streamflow signal at the Whitemouth gauge, despite the statistically significant anthropogenic precipitation signal at the nearby station, echoes this same pattern.

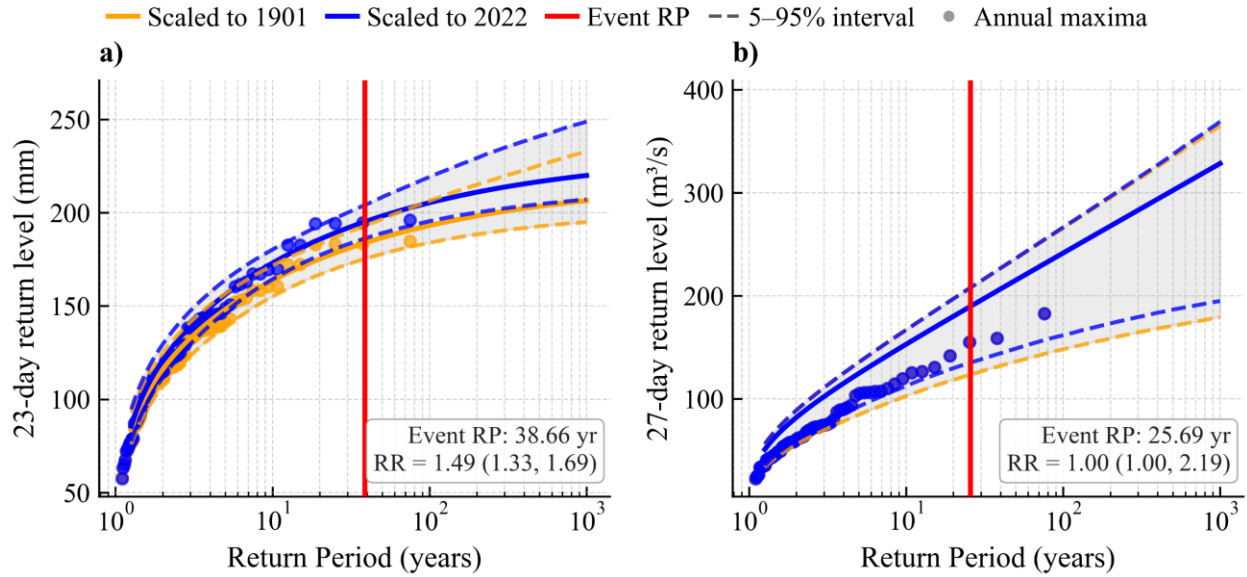


Figure 9: Return-level curves and corresponding risk ratios (RR) for annual maximum precipitation at 49.02°N, 95.60°W (a) and streamflow at the Whitemouth River gauge 05PH003, 49.94°N, 95.96°W (b), derived using the scaling-based method. Blue curves represent the factual climate (with anthropogenic forcing) and orange curves represent the counterfactual climate (natural forcing only). Solid lines denote fitted GEV return-level curves, dashed lines indicate the 5th–95th percentile uncertainty bounds (90% interval), and the vertical red line marks the return period of the 2022 event estimated from the GEV fit to the observed record. Reported RR values correspond to the probability ratio of the event magnitude under factual versus counterfactual conditions.

4.4.2 Time-Slice PEA Analysis

Applying the time-slice attribution method, in which the observational record at each station is partitioned into early and late climate slices separated by a detected change-point year, supports the anthropogenic signature noted in the scaling analysis and yields a modestly larger fraction of significant detections for both variables (Figure 10). The Pettitt-based change points cluster around the mid-1980s for both precipitation (median 1986, range 1965 to 1994) and streamflow (median 1986, range 1969 to 1995), consistent with the late-twentieth-century

acceleration of observed warming over southern Canada. Similar to the scaling results, the majority of stations exhibit $RR \geq 1$ at the point-estimate level, with 14 of the 19 precipitation stations (about 74%) and 20 of the 28 streamflow gauges (about 71%) exceeding unity (Figure 10, center column). At the 90% confidence level, 8 of the 19 precipitation stations and 6 of the 28 streamflow gauges produce confidence intervals lying entirely above unity, representing one additional significant detection in each variable relative to the scaling results. Two precipitation stations (49.56°N, 99.07°W and 49.40°N, 98.32°W) and one streamflow gauge (53.16°N, 99.35°W) yield confidence intervals lying entirely below unity, indicative of dominant natural variability. A further feature of the time-slice results is that the magnitudes of the central RR estimates are considerably larger than those produced by the scaling method, particularly in the eastern sub-region where multiple stations return RR values exceeding 10. This amplification is an expected consequence of comparing two independently fitted GEV distributions, where shifts in the tail between the early and late slices translate into large probability ratios at the event magnitude, accompanied by wider confidence bounds.

The east-west spatial gradient identified in the scaling analysis is preserved and further sharpened under the time-slice method (Figure 10), with a weak central-south signal and stronger anthropogenic influence in the eastern sub-region. In the central-south sub-region, point RR estimates remain near unity for both variables (median precipitation RR of 1.04 and median streamflow RR of 1.20), and no stations reach 90% significance for anthropogenic influence; one station instead yields a significant natural-variability signal. The central-south null result is not a consequence of insufficient statistical power: the Pettitt test detects a significant change point at seven of the eight central-south precipitation stations ($p < 0.05$), so the partitioning at this sub-region is constrained by data-driven breakpoints rather than the median-split fallback. However,

six of these seven stations are empirically drier in the post-change-point period than in the pre-change-point period, and the sub-regional median empirical shift in the GEV location parameter is -13.43 mm, compared with $+21.06$ mm at the eastern stations and $+6.42$ mm at the western stations (Appendix Figure C7). The detected breakpoints therefore reflect non-monotonic shifts associated with internal climate variability rather than progressive intensification under warming. In eastern Manitoba, the anthropogenic signal is considerably stronger than in the scaling analysis: all 7 precipitation stations exhibit point $RR \geq 1$ and 6 of the 7 are significant at the 90% level, with a sub-regional median RR of 6.26, while all 5 co-located streamflow gauges likewise exhibit point $RR \geq 1$ and return a sub-regional median of 6.69. In western Manitoba, 2 of the 3 precipitation stations reach 90% significance and 3 of the 8 streamflow gauges in the combined western and north-central area also cross the significance threshold, indicating that the modest anthropogenic tendency noted for this sub-region is consistent with the scaling-based results.

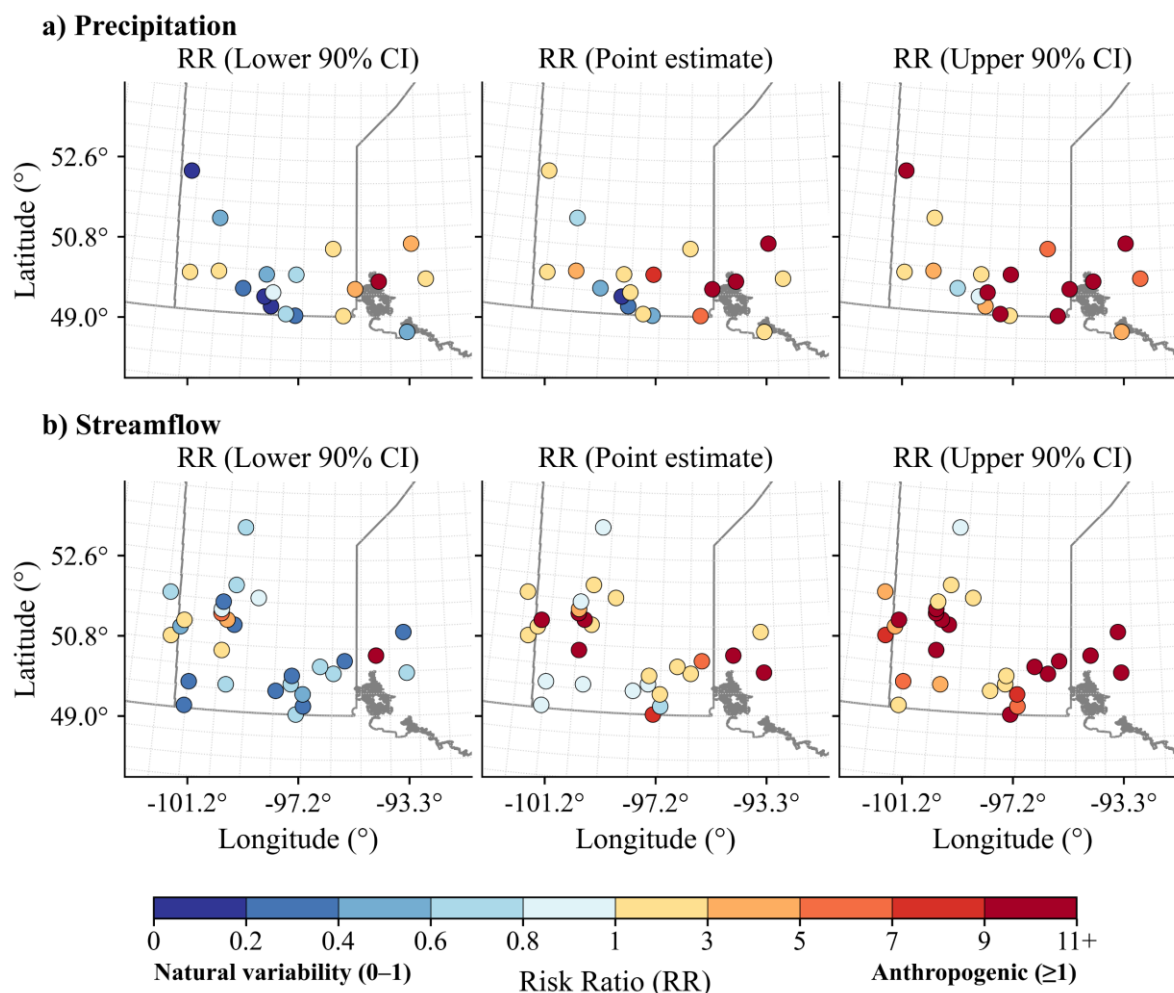


Figure 10: Spatial distribution of station-level risk ratios (RR) derived from the time-slice attribution method. Panels show results for (a) precipitation (19 stations) and (b) streamflow (28 gauges), with change-point years at each site identified using the Pettitt test and a median-split fallback. For each variable, the left, center, and right columns display the lower bound of the 90% confidence interval, the point estimate, and the upper bound of the 90% confidence interval, respectively. Colors denote RR magnitude, with blue shades ($RR < 1$) indicating dominance of natural variability and warm shades ($RR \geq 1$) indicating a central estimate of anthropogenic influence.

Station-scale return-level curves for a representative precipitation station and streamflow gauge further illustrate the time-slice behavior (Figure 11). At the precipitation station at 50.52°N , 95.98°W in eastern Manitoba, the Pettitt test identifies a change point at 1993, partitioning the record into an early slice (1950 to 1993) and a late slice (1994 to 2022). Comparison of the two

GEV fits for the 29-day accumulation of 142.7 mm observed in 2022 yields a point-estimate RR of 1.69 with a 90% confidence interval of [1.42, 6.63] (Figure 11a). The lower bound exceeding unity indicates a statistically significant increase in event probability, with the observed magnitude approximately 69% more likely under the post-1993 climate, corresponding to a fraction of attributable risk of about 0.41. The wide upper bound reflects limited constraint on the upper tail of the late-slice GEV fit given the 29-year post-change-point record. At the streamflow gauge at 50.35°N, 99.91°W in western Manitoba, the Pettitt-informed partitioning places the change point at 1990, producing balanced early (1959 to 1990) and late (1991 to 2022) slices of 32 years each. The 2022 peak 14-day mean flow of $106.5 \text{ m}^3 \text{ s}^{-1}$ yields a point-estimate RR of 14.32 with a 90% confidence interval whose upper bound is extremely large (Figure 11b). The large central RR reflects the pronounced upward displacement of the late-slice GEV return-level curve relative to the early-slice curve at this gauge, while the extreme upper confidence bound is a direct result of the bootstrap sampling of a heavy-tailed GEV shape parameter under a small post-change-point sample; despite this, the lower bound remains above unity, which places the result within the statistically significant anthropogenic category. Together, these contrasting precipitation and streamflow examples highlight that the time-slice method identifies anthropogenic influence across a broader portion of the study area than the scaling method, although with substantially larger uncertainty ranges on individual RR estimates.

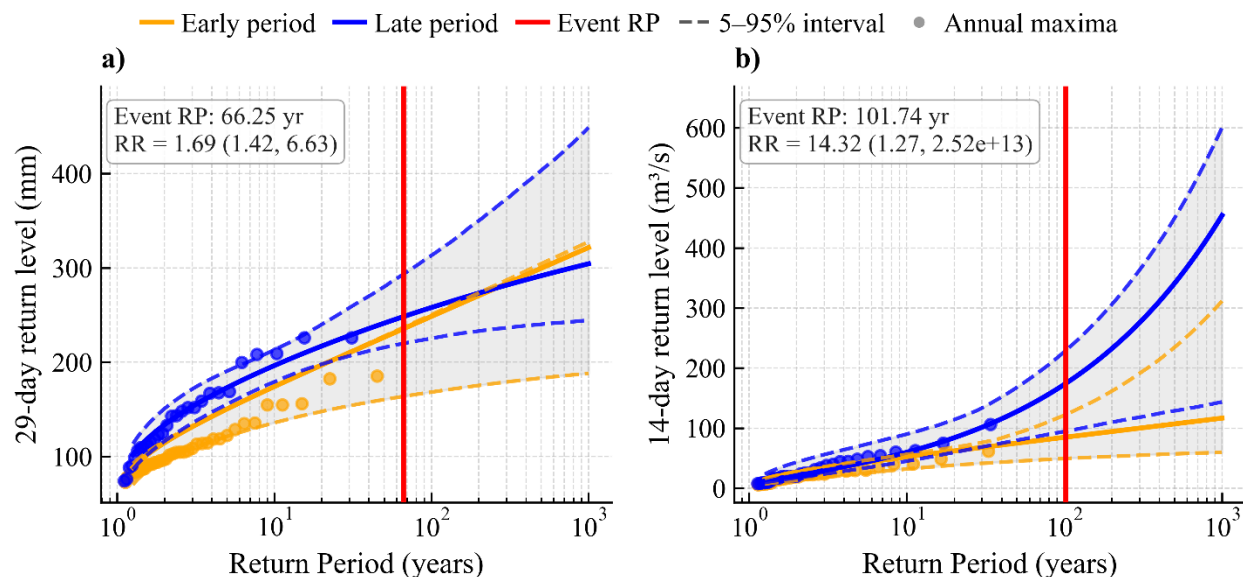


Figure 11: Return-level curves and corresponding risk ratios (RR) for annual maximum 29-day precipitation at $50.52^{\circ}N$, $95.98^{\circ}W$ (a) and annual maximum 14-day mean flow at the streamflow gauge at $50.35^{\circ}N$, $99.91^{\circ}W$ (b), derived using the time-slice attribution method. Orange curves represent the early climate slice (pre-change-point) and blue curves represent the late climate slice (post-change-point). Solid lines denote fitted GEV return-level curves, dashed lines indicate the 5th–95th percentile uncertainty bounds (90% interval), and the vertical red line marks the return period of the 2022 event estimated from the GEV fit to the observed record. Reported RR values correspond to the probability ratio of the event magnitude under the late versus early climate slice.

4.4.3 GCM-Based PEA Analysis for Precipitation

The GCM-based PEA analysis was performed using four CMIP6 models (BCC-CSM2-MR, CNRM-CM6-1, IPSL-CM6A-LR, and NorESM2-LM; per-model results are shown in Appendix Figure C4) participating in the Detection and Attribution Model Intercomparison Project (DAMIP; Gillett et al., 2016), which provides paired historical (all-forcing) and historical-natural (natural-only) simulations required to construct the factual and counterfactual climates. For precipitation, the analysis was applied at all 19 stations retained for this method; for streamflow, the analysis was restricted to a single representative gauge because hydrological simulations driven

by CMIP6 forcings were generated only at that location in this study. In the multi-model ensemble mean, 13 of the 19 precipitation stations (about 68%) exhibit point RR values at or above unity (Figure 12, center column), a proportion comparable to the observation-based scaling and time-slice methods. At the 90% confidence level, 5 stations produce confidence intervals lying entirely above unity, and no station yields a confidence interval entirely below unity. The overall amplification of the ensemble-mean response, with a study-area median point RR of 1.53, is consistent with the broader literature documenting detectable anthropogenic influence on Canadian extreme precipitation when aggregated across multiple CMIP6 models (Dong et al., 2021).

The spatial distribution of ensemble-mean RR values preserves the same east-west contrast identified in the scaling and time-slice analyses, with a sharper amplification of the off-Prairie signals (Figure 12). The central-south sub-region is anchored in the natural-variability regime, with 3 of 8 stations producing a point RR at or above unity, none reaching 90% significance, and a sub-regional median RR of 0.76. The corresponding sub-regional forced response in the GEV location parameter is +5.08 mm at the central-south stations, the smallest forced response of any sub-region, compared with +7.94 mm at the eastern stations and +7.19 mm at the western stations (Appendix Figure C7). The convergence of the smallest CMIP6 forced response with the negative scaling coefficients (Section 4.4.1) and the empirically drier post-change-point period (Section 4.4.2) supports the interpretation that the central-south sub-region sits in a regime of limited anthropogenic intensification of 29-day extremes. Eastern Manitoba retains the strongest anthropogenic tendency, with 5 of 6 stations exhibiting point $RR \geq 1$, 2 of 6 reaching 90% significance, and a sub-regional median RR of 2.79. Western Manitoba shows a uniformly anthropogenic-leaning signal under the ensemble mean: all 3 stations exhibit point $RR \geq 1$ and all

3 reach 90% significance, with a sub-regional median RR of 3.27. The two northern stations likewise produce point $RR \geq 1$ with a median of 9.95, although neither reaches 90% significance owing to wide bootstrap confidence intervals.

A substantial inter-model spread is noted across the four individual CMIP6 models (Appendix Figure C4), which is informative about the robustness of the ensemble-mean pattern. BCC-CSM2-MR produces a uniformly anthropogenic-leaning response, with 17 of 19 stations at or above unity, the most pronounced signal concentrated in eastern Manitoba (sub-regional median RR of 4.59, with 3 of 6 stations significant at the 90% level), and a coherent anthropogenic tendency over the central-south (median RR of 1.78) and western (median RR of 1.29) sub-regions; only the two northern stations depart into the natural-variability regime. CNRM-CM6-1 produces the weakest and most mixed response of the four models, with 9 of 19 stations at point $RR \geq 1$, only one station reaching 90% significance for anthropogenic influence and one for natural variability, and a study-area median RR of 0.90; its anthropogenic tendency is confined largely to the central-south-to-eastern belt (median RR of 1.13 and 1.09, respectively) while the western and northern portions sit in the natural-variability regime. IPSL-CM6A-LR places the natural-variability signal in the central-south (median RR of 0.85) while producing a strong anthropogenic signal in the eastern (median RR of 2.05, with 3 of 6 stations significant) and western (median RR of 1.87) sub-regions, along with a notably amplified signal at the northern stations (median RR of 15.82). NorESM2-LM produces the strongest anthropogenic-leaning response of the four models, with 16 of 19 stations at point $RR \geq 1$, the largest fraction of significant detections (10 of 19), and pronounced sub-regional medians over the western (10.78, all 3 stations significant), northern (55.6), and central-south (6.15, with 4 of 8 stations significant) portions of the study area, and a moderate signal over the east (median RR of 2.62). The wide disagreement among the four models

is consistent with the documented sensitivity of CMIP6 precipitation responses to each model's representation of aerosol forcing, convective parameterization, and large-scale circulation biases (Dong et al., 2021; Shiogama et al., 2022); because aerosol forcing in particular remains one of the largest sources of inter-model spread in the historical period, individual-model RR estimates should be interpreted with caution and the ensemble-mean response should be regarded as the more reliable indicator of the externally forced signal (Dong et al., 2021; Gillett et al., 2016).

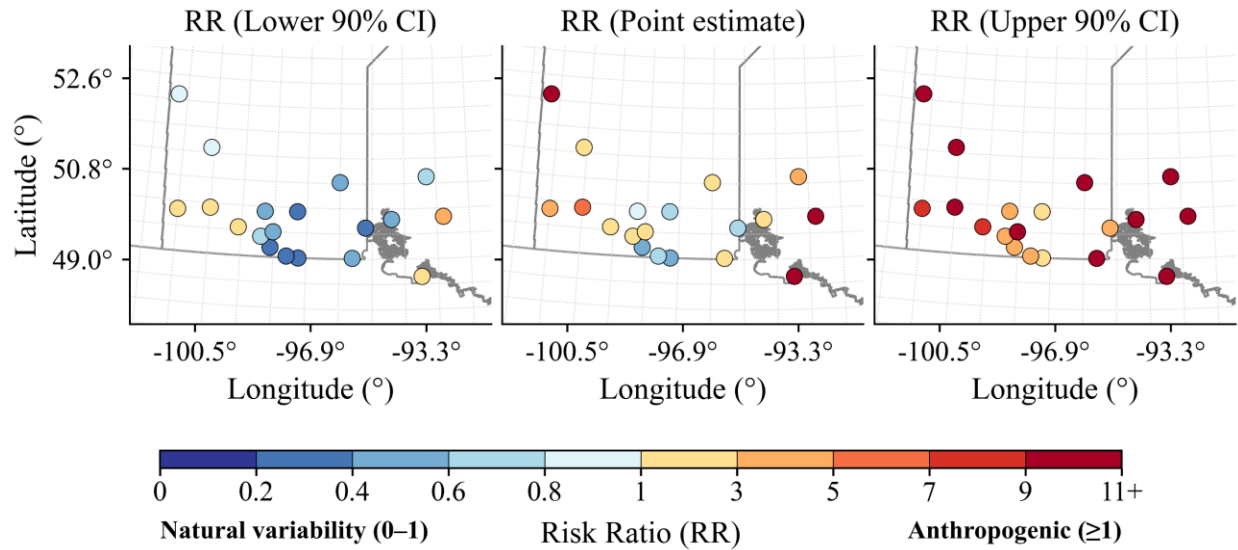


Figure 12: Spatial distribution of station-level risk ratios (RR) derived from the GCM-based attribution method, using the four-model CMIP6 ensemble mean. The left, center, and right columns display the lower bound of the 90% confidence interval, the point estimate, and the upper bound of the 90% confidence interval, respectively. Colors denote RR magnitude, with blue shades ($RR < 1$) indicating dominance of natural variability and warm shades ($RR \geq 1$) indicating a central estimate of anthropogenic influence.

Station-scale return-level curves for the ensemble-mean fit at a representative eastern station (49.83°N , 92.80°W) further illustrate the GCM-based attribution behavior (Figure 13). For the 29-day precipitation event at this station, the 2022 magnitude of 146.5 mm yields a point-estimate RR of 10.35 with a 90% confidence interval of [3.63, 117.49], corresponding to an event approximately 10 times more likely under the historical all-forcing climate than under the

historical-natural climate and a fraction of attributable risk of about 0.90. Because the lower bound of the confidence interval substantially exceeds unity, the result is statistically significant at the 90% level. The pronounced upward displacement of the historical GEV return-level curve relative to the historical-natural curve is evident across the full range of return periods, consistent with the direction of the signal at this station under the scaling and time-slice methods and reinforcing the interpretation that eastern Manitoba shows the most persistent and strongly amplified anthropogenic influence across all three attribution frameworks examined in this study.

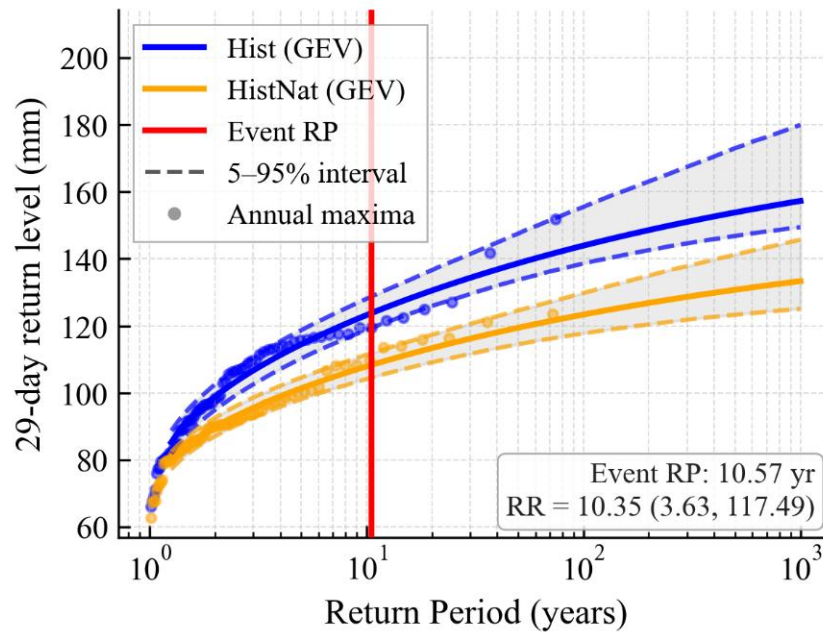


Figure 13: Return-level curves and corresponding risk ratio (RR) for annual maximum 29-day precipitation at the representative station at 49.83°N, 92.80°W, derived using the GCM-based attribution method with the four-model CMIP6 ensemble mean. The blue curve represents the historical all-forcing (factual) climate, and the orange curve represents the historical-natural (counterfactual) climate. Solid lines denote fitted GEV return-level curves, dashed lines indicate the 5th–95th percentile uncertainty bounds (90% interval), and the vertical red line marks the return period of the 2022 event estimated from the GEV fit to the observed record.

4.4.4 *Performance of the LSTM and Raven Streamflow Simulation*

The HistNat scenario streamflow for the four GCMs was obtained using two hydrological models, LSTM and Raven, at the Whitemouth River gauge (49.94°N, 95.96°W). Both models were calibrated, tested, and validated before obtaining the streamflow for the HistNat scenario (training diagnostics and flow duration curves in Appendix Figures B1 and B2), and both reproduce the observed hydrological variability at this gauge with sufficient skill to support the subsequent GCM-based streamflow attribution (Figure 14). The LSTM model delivers the stronger overall performance, achieving $NSE = 0.91$ and $KGE = 0.92$ during training (1950 to 2010) with an RMSE of $8.39 \text{ m}^3 \text{ s}^{-1}$, and maintaining robust skill on the held-out validation (2011 to 2016; $NSE = 0.78$, $KGE = 0.77$, $RMSE = 13.13 \text{ m}^3 \text{ s}^{-1}$) and test (2017 to 2023; $NSE = 0.75$, $KGE = 0.70$, $RMSE = 14.60 \text{ m}^3 \text{ s}^{-1}$) periods (Figure 14a). RMSE increases from training to validation and test, but NSE and KGE remain within a range that supports use of the model for the subsequent attribution analysis.

The Raven process-based model exhibits moderate skill at the same gauge (Figure 14b). Performance over the calibration period (1955 to 2010) is $NSE = 0.53$, $KGE = 0.66$, $RMSE = 19.00 \text{ m}^3 \text{ s}^{-1}$, and skill improves on the independent validation window (2011 to 2023) to $NSE = 0.69$, $KGE = 0.68$, $RMSE = 16.05 \text{ m}^3 \text{ s}^{-1}$. The time-series comparison shows that Raven reproduces the seasonal evolution of streamflow, including the onset and recession of the spring snowmelt freshet and the transition to baseflow through summer and autumn (Figure 14b). Deviations between observed and simulated series are more pronounced at higher flows, reflecting the well-recognized difficulty of reproducing extreme runoff generation in flat, prairie-shield transitional catchments.

Flow-duration curves (Appendix Figure B2) reinforce these results. Both models reproduce the observed flow distribution closely during their training or calibration periods and track the bulk of the mid-range (20 to 70%) reasonably well in validation and test, with the largest discrepancies occurring in the high-flow ($p < 5\%$) and low-flow ($p \geq 70\%$) tails and a modest mid-range bias in the LSTM validation and test windows. Overall, both models demonstrate stable skill outside their calibration periods, and the LSTM in particular provides the fidelity required to translate CMIP6 precipitation forcings into physically plausible streamflow time series for the GCM-based attribution of the 2022 event.

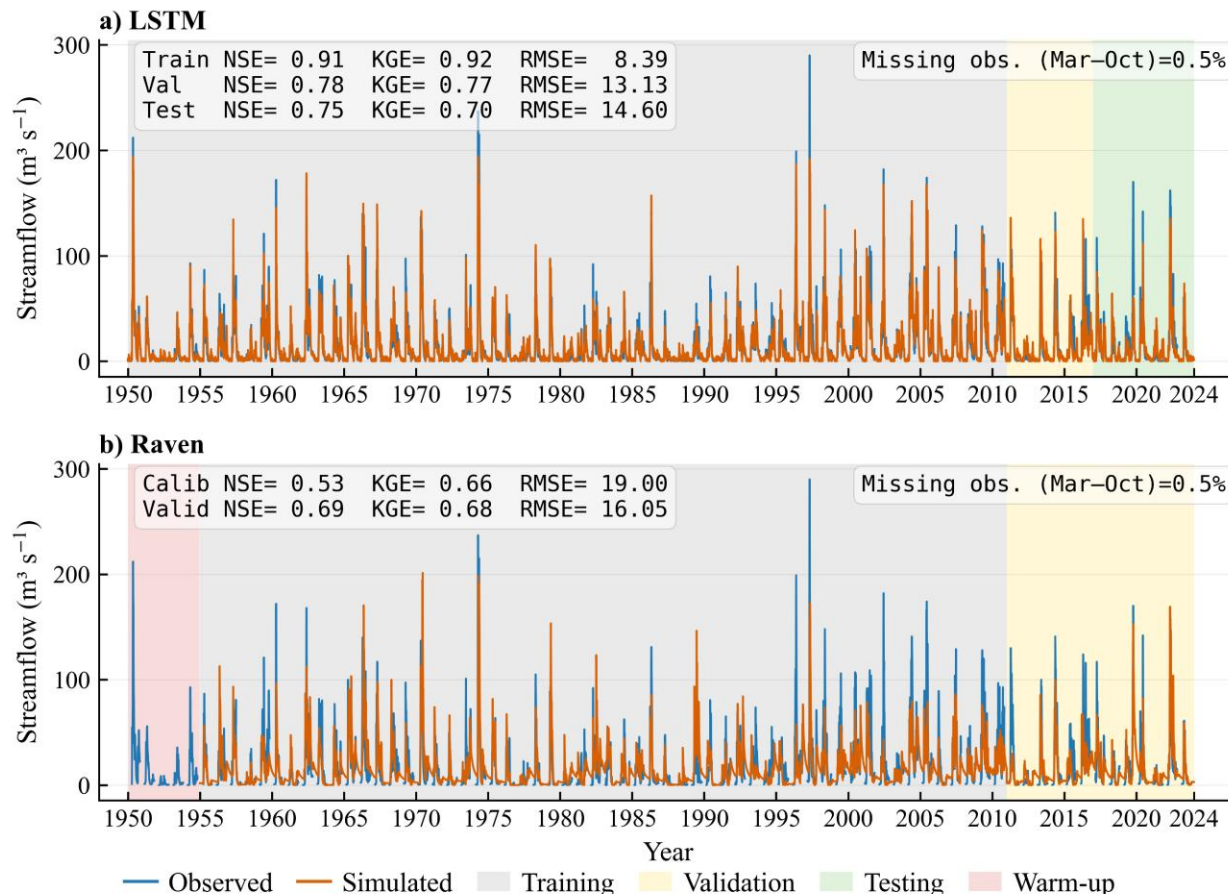


Figure 14: Observed (blue) and simulated (orange) daily streamflow at the Whitemouth River gauge (05PH003) for (a) the LSTM data-driven model, evaluated over training (1950 to 2010), validation (2011 to 2016), and test (2017 to 2023) periods, and (b) the Raven process-based model, evaluated over calibration (1955 to 2010) and validation (2011 to 2023) periods. Reported metrics are Nash-Sutcliffe efficiency (NSE), Kling-Gupta efficiency (KGE), and root-mean-square error (RMSE, $\text{m}^3 \text{s}^{-1}$). Shaded backgrounds denote the evaluation windows: training/calibration (gray), validation (yellow), testing (green), and warm-up (pink).

4.4.5 GCM-Based PEA Analysis for Streamflow

The historical and historical-natural CMIP6 precipitation and temperature time series from the four GCMs (BCC, CNRM, IPSL, NorESM2) were used as inputs to both the data-driven LSTM model and the process-based Raven model to simulate the 27-day mean flows at the Whitemouth

River gauge. For the 2022 event, defined as a 27-day mean flow of $141.6 \text{ m}^3 \text{ s}^{-1}$ (empirical return period of about 25.7 yr), the two hydrological frameworks produce multi-model ensemble risk ratios that differ in magnitude and approach to unity.

The LSTM-based ensemble indicates an RR of 1.45 [0.65, 2.66], indicating that the 2022 flood was slightly more likely under the factual (all-forcing) climate relative to the counterfactual (HistNat) climate (Figure 15a). Individual GCM contributions show variability, with BCC yielding 1.48 [0.62, 4.29], CNRM 1.02 [0.40, 4.26], IPSL 2.66 [0.98, 7.83], and NorESM2 0.65 [0.35, 1.76], reflecting the sensitivity of the ensemble to higher-RR models such as IPSL.

The Raven-based simulations produce a higher ensemble RR of 3.14 [0.83, 8.48] (Figure 15b), with per-model RRs of 0.95 (BCC), 3.42 (CNRM), 4.43 (IPSL), and 1.76 (NorESM2). The elevated RR in Raven highlights the influence of hydrological model structure, including snowmelt parameterization, soil-moisture dynamics, and routing, on the attribution signal. Across both LSTM and Raven, the attribution direction is model-dependent, with some models indicating enhanced likelihood under anthropogenic forcing ($\text{RR} > 1$) and others a weaker signal ($\text{RR} < 1$). None of the individual LSTM RRs reaches 90% statistical significance, whereas Raven shows wider model spread with certain runs approaching significance. Per-model return-level curves are provided in Appendix Figures C5 (LSTM) and C6 (Raven).

The lower RR obtained from the LSTM model relative to Raven likely reflects structural differences between the two frameworks. The LSTM is a data-driven model that learns statistical relationships from historical flow records and may underpredict extreme flows if such events are underrepresented in the training period. In contrast, Raven is a process-based model that explicitly simulates snow accumulation and melt, soil-moisture storage, and routing, allowing for a physically consistent propagation of precipitation extremes into streamflow peaks, which explains

the higher RR in Raven. Taken together, the ensemble RRs from LSTM and Raven show anthropogenic influence with a large model-to-model spread and overlapping 90% confidence intervals, indicating no statistically robust anthropogenic signal at this gauge. This contrasts with the precipitation-based attribution (Sections 4.4.1 to 4.4.3), which shows a detectable anthropogenic signal, and reflects the integration and buffering effects of snowpack, soil moisture, and hydrological routing on streamflow (Teufel et al., 2017).

These GCM-based results align with the scaling method (Section 4.4.1), which yields $RR = 1.00$ (CI [1.00, 2.19]), and the time-slice method, which produces a point estimate of 2.52 with a CI spanning unity. Across all methods, the central RR estimates range from unity to moderately above it, and none exclude the null hypothesis, reinforcing the interpretation that the 2022 Whitemouth River flood cannot be confidently attributed to anthropogenic forcing at this location, despite strong anthropogenic signals in co-located precipitation.

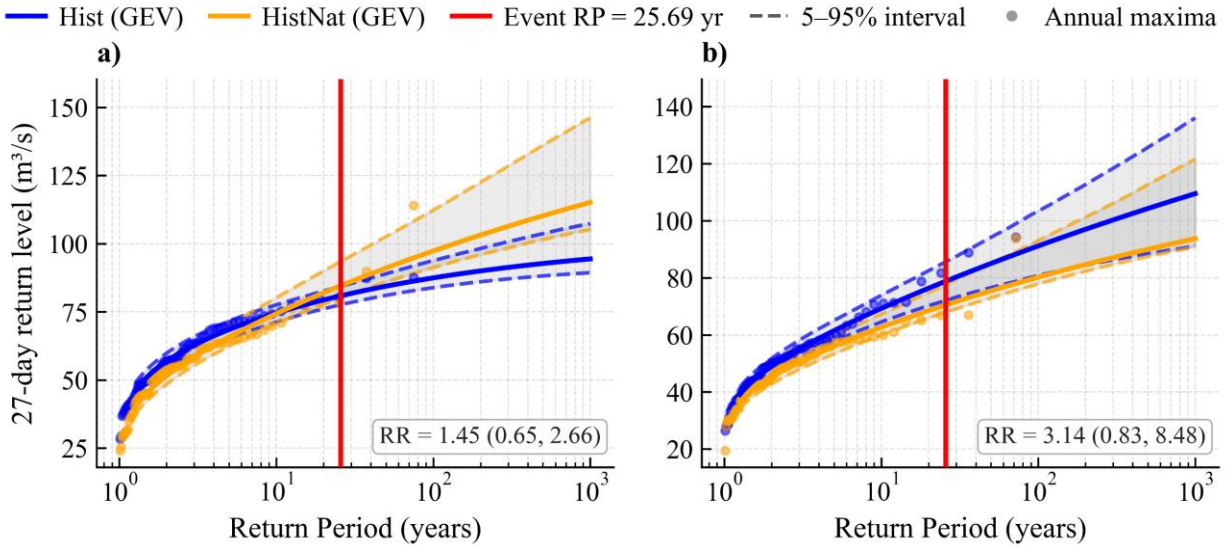


Figure 15: Ensemble-mean return-level curves and corresponding risk ratios (RR) for annual maximum 27-day mean flow at the Whitemouth River gauge (49.94°N, 95.96°W) from the GCM-based attribution method using (a) the LSTM data-driven hydrological model and (b) the Raven process-based hydrological model. Blue curves represent the historical all-forcing (factual) climate; orange curves represent the historical-natural (counterfactual) climate. Solid lines denote fitted GEV return-level curves, dashed lines the 5th–95th percentile bootstrap bounds, and circles show the annual maxima of the Hist and HistNat simulations at empirical plotting positions. The vertical red line marks the return period of the 2022 event (25.7 years) estimated from the GEV fit to the observed record.

4.4.6 Bivariate (Compound) Attribution Results

The 2022 cool-spring plus high-snow event was analyzed using copula-based bivariate attribution at Indian Bay (49.62°N, 95.20°W), located in the eastern Lake of the Woods region that was directly relevant to the 2022 flood-generating conditions in the Winnipeg River basin. Indian Bay was selected to represent the joint snow-temperature preconditioning associated with the Spring 2022 event because it captures the eastern sub-region where antecedent snow accumulation and cool spring temperatures contributed to delayed snowmelt and elevated flood potential.

At Indian Bay, the Pettitt test did not detect a significant change point in the November to March peak snow-depth series ($p = 0.763$). Therefore, the default 1985 split-year was retained, separating the record into an early period (≤ 1985 ; $n = 36$) and a late period (> 1985 ; $n = 38$). For both periods, the best-fitting marginal distributions and copula families were selected using AIC-based ranking and a rank-sum criterion. In the early period, the April mean temperature marginal was best represented by a normal distribution, the snow-accumulation marginal by a Weibull distribution, and the dependence structure by a Frank copula. In the late period, the temperature marginal shifted to a Student-t distribution, the snow marginal shifted to a gamma distribution, and the dependence structure transitioned to a Gaussian copula.

This shift from a Frank copula in the early period to a Gaussian copula in the late period suggests a change in the dependence structure between cool spring temperatures and high peak snow depth at Indian Bay. The late-period structure indicates a modest tightening of the joint distribution, meaning that cool-spring and high-snow conditions became slightly more concordant after the mid-1980s. The fitted marginal distributions are shown in Figure 16a, while the corresponding risk-ratio distributions are summarized using box plots in Figure 16b. The early- and late-period joint distributions are shown in Figure 16c and Figure 16d, respectively.

At Indian Bay, the joint exceedance probability of 2022-like compound conditions increased from 0.015 in the early period (≤ 1985) to 0.017 in the late period (> 1985). This produced a compound risk ratio of 1.11, with a 90% confidence interval of 0.13 to 5.12 and a FAR of 0.10. The central estimate is marginally above unity, but the confidence interval spans a wide range around the no-change line, indicating that the joint-probability increase is not statistically distinguishable from no change at the 90% confidence level.

The marginal components straddle unity in opposite directions. The temperature marginal risk ratio is 0.96, with a 90% confidence interval of 0.38 to 2.39, indicating no statistically significant change in the probability of April mean temperatures at or below the 2022 value. The snow marginal risk ratio is 0.74, with a 90% confidence interval of 0.18 to 2.03, suggesting a modest decline in the probability of November to March peak snow depth at or above the 2022 value, consistent with the general warming-driven reduction in snow accumulation seen across cold-region records, though this decline is also not statistically significant. Neither marginal change alone is robustly attributable; the slight rise in compound risk reflects the dependence-structure tightening discussed above rather than a shift in either marginal.

The 2022 event return periods estimated from the late-period marginals are 12.9 years for the temperature component and 16.0 years for the snow component. Both variables were therefore individually moderately exceptional, placing the 2022 event well into the upper-right corner of the joint distribution (Figure 16d). The compound joint probability of approximately 0.017 corresponds to a joint return period of roughly 60 years in the late period, substantially rarer than either variable in isolation. Although the inter-period change in joint probability is small relative to the bootstrap uncertainty ($RR = 1.11$), the compound framing reveals that the 2022 preconditioning was driven by the co-occurrence of a moderately cool spring with a moderately deep peak snowpack, neither of which would be unusual on its own but whose joint occurrence is rare in either period.

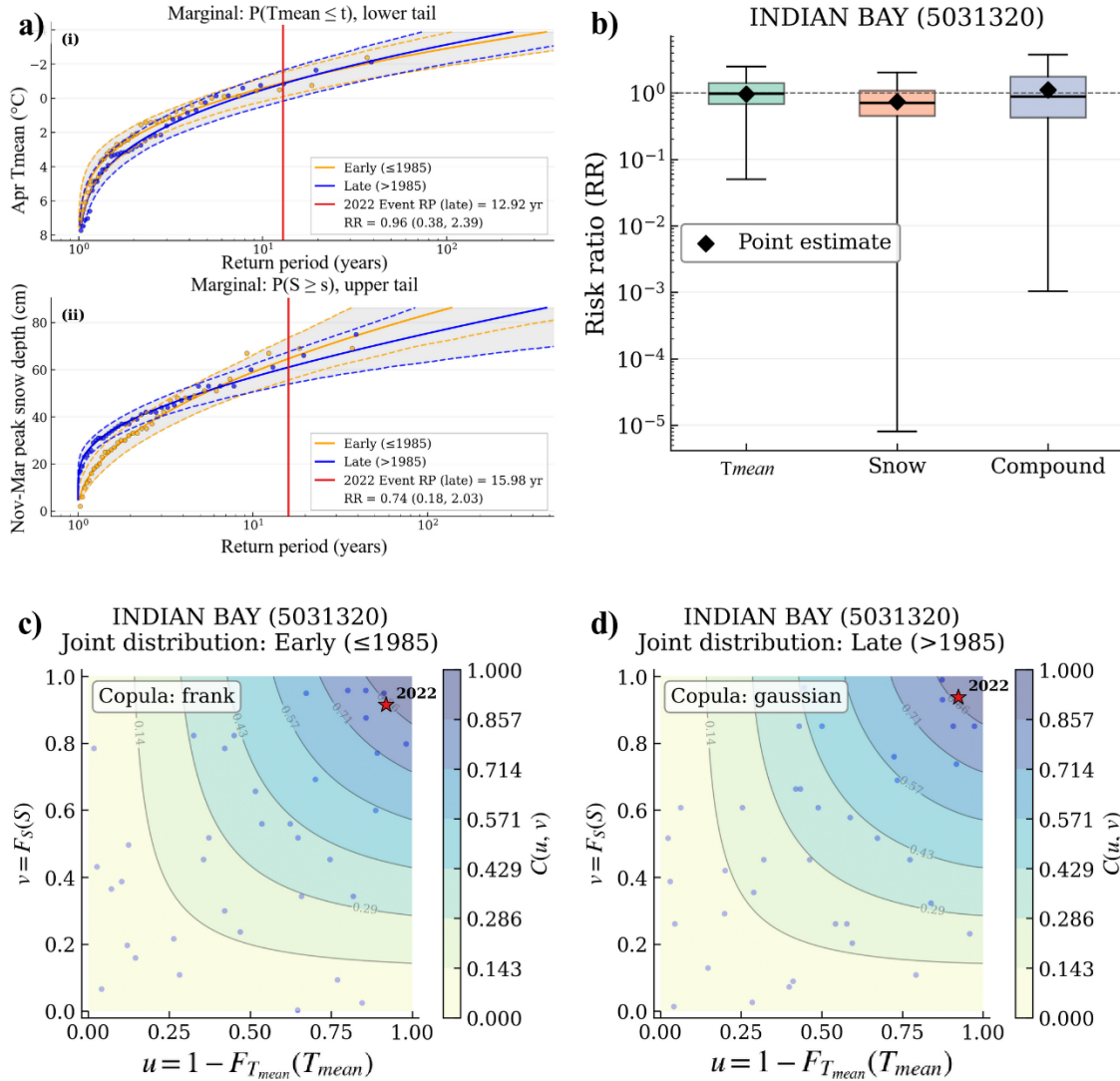


Figure 16: Bivariate compound attribution results for cool-spring and high-snow preconditioning at Indian Bay (5031320). Panel (a) shows the fitted marginal return-level curves for the April mean temperature lower tail, $P(T_{mean} \leq t)$, and the November to March peak snow-depth upper tail, $P(S \geq s)$, for the early and late periods. The 2022 event value is marked by the vertical red line, and uncertainty bounds are shown using dashed lines. Panel (b) shows box plots of the bootstrap distributions of the marginal temperature, marginal snow, and compound risk ratios, with the point estimate marked separately. Panels (c) and (d) show the fitted joint distributions in copula uniform space, where $u = 1 - F_{T_{mean}}(T_{mean})$ represents cool-spring conditions and $v = F_S(S)$ represents high antecedent snow accumulation. Panel (c) shows the early period and panel (d) shows the late period, with the 2022 event location marked by a red star.

4.5 *Discussion and Conclusions*

This chapter evaluated the role of anthropogenic climate change in the Spring 2022 precipitation and flooding event in southern Manitoba and northwestern Ontario using a probabilistic event attribution (PEA) framework. The application of three PEA frameworks reveals a consistent qualitative picture but substantial quantitative divergence. Across precipitation and streamflow, central risk-ratio estimates generally suggest an increased likelihood of the Spring 2022 extremes under present-day conditions. The three methods agree on the direction of the signal but differ substantially in its magnitude, a pattern that reinforces the value of multi-method attribution for complex hydroclimatic extremes.

Differences across methods reflect their underlying assumptions and sensitivities. The scaling approach isolates the thermodynamic effects of warming using observations but cannot account for circulation changes or land-surface feedback. The time-slice method implicitly incorporates all sources of long-term change and therefore yields the largest central estimates, but at the cost of large statistical uncertainty due to limited sample sizes and sensitivity to slice definition. The GCM-based framework explicitly contrasts factual and counterfactual climates but is constrained by coarse spatial resolution and ensemble spread, particularly for multi-week precipitation extremes. Taken together, the three approaches draw on different information and rest on different assumptions, reinforcing confidence in the qualitative conclusions while underscoring the limits to precise quantification. Furthermore, the comparison between the data-driven LSTM and process-based Raven hydrological models provides an important consistency check on the streamflow attribution results. Despite their fundamentally different structures and assumptions, both models yield central risk-ratio estimates that lie in the anthropogenic-leaning regime yet span unity at the 90% level for the Spring 2022 flood at the Whitemouth River, with overlapping

uncertainty ranges. This agreement across structurally distinct hydrological models suggests that the absence of a statistically detectable anthropogenic signal in flood magnitude is unlikely to be an artifact of model choice. Instead, it reflects dominant physical and climatic uncertainties in runoff generation, including the influence of antecedent snowpack, frozen-soil conditions, basin storage, and internal climate variability, which together limit the detectability of forced changes in extreme streamflow at the basin scale.

4.5.1 Diagnostic Synthesis of the Regional Attribution Signal

The east-to-west gradient in risk ratios across the scaling, time-slice, and GCM-based methods (Sections 4.4.1 to 4.4.3) raises a question: does the absence of a detectable anthropogenic signal over the central-south sub-region reflect a true absence of an anthropogenic signal, or does it simply reflect weak statistical power at individual stations? To answer this, we extracted the shift in the GEV location parameter at every station, computed separately under each of the three methods, and aggregated the values by sub-region (Appendix Figure C7). For the scaling method, the shift is the difference between the location parameter fitted at the 2022 climate state and that fitted at the 1901 state; for the time-slice method, it is the empirical difference between the late-period and early-period location parameters; and for the GCM method, it is the difference between the all-forcing and natural-only location parameters in the multi-model ensemble mean. Under all three methods, a positive value means that 29-day precipitation extremes are larger in the warmer climate state.

Across the eight central-south stations, all three diagnostics return non-positive median shifts. The fitted scaling coefficient has a sub-regional median of -1.10 , giving a scaling-method shift of -1.93 mm between the 1901 and 2022 states. The empirical time-slice shift has a sub-regional median of -13.43 mm, with six of eight stations showing a non-positive value, and the

CMIP6 anthropogenic signal is +5.08 mm, the smallest of any sub-region. The eastern stations sit in a different regime, with a median empirical time-slice shift of +21.06 mm and a CMIP6 anthropogenic signal of +7.94 mm, more than 1.5 times the central-south value. The western stations fall between these two regimes, with a positive median scaling coefficient (+0.61), a positive empirical shift (+6.42 mm), and a positive CMIP6 anthropogenic signal (+7.19 mm). As shown in Section 4.4.2, the central-south breakpoints are data-driven but track internal variability rather than progressive intensification.

To confirm that this regional contrast represents a real climate signal and is not produced by the bootstrap procedure, we computed sub-regional trends over 1950 to 2022 in spring precipitation, spring mean temperature, spring evapotranspiration, snow-season precipitation, and spring peak streamflow at each station, using the Mann-Kendall test with Theil-Sen slope estimation and 90% bootstrap confidence intervals (Appendix Figure C8). The most informative driver is the April to May precipitation total, which corresponds directly to the spring rainfall that, together with the antecedent snowpack, generated the 2022 flood. The median Theil-Sen slope is +0.27 mm/decade in the central-south (positive at five of eight stations), +2.00 mm/decade in the east (positive at six of six stations), and +0.59 mm/decade in the west (positive at three of four stations). A two-sided Mann-Whitney U test shows that the central-south and eastern slope distributions differ significantly ($U = 5$, $p = 0.013$, rank-biserial $r = +0.79$), corresponding to a large effect size. The same direction of contrast is evident in April to May mean temperature, where the eastern stations warm more rapidly than the central-south stations. In contrast, sub-regional trends in March to May evapotranspiration (+5.75 mm/decade central-south, +6.85 mm/decade east; $p = 0.41$) and in November to February snow-season precipitation (−5.36 mm/decade central-south, −5.95 mm/decade east; $p = 1.00$) are statistically indistinguishable, ruling out spring

evapotranspiration and winter snow accumulation as drivers of the contrast in the GEV-based signal. The streamflow response mirrors the precipitation contrast: the median April to May peak-flow trend is $+0.24 \text{ m}^3 \text{ s}^{-1}$ per decade at the four central-south gauges and $+1.13 \text{ m}^3 \text{ s}^{-1}$ per decade at the five eastern gauges.

These driver trends are consistent with recent observations of changing spring precipitation regimes across northwestern Ontario and southeastern Manitoba, although the underlying mechanisms remain unsettled. Recent large-scale analyses disagree on whether spring extratropical cyclone activity has changed: Li et al. (2025) report an increase in March to May cyclone activity over the Northern Hemisphere oceans, whereas Lynne & Dai (2025) find no significant change for April in ERA5. Regional warming over the Canadian Prairies and the lengthening ice-free season at Lake of the Woods and the upstream Lake Winnipeg sub-system are also relevant (Bonsal et al., 2019; Mudryk et al., 2015; Vincent et al., 2018). Over the Red River Valley, by contrast, spring precipitation is sourced largely from Gulf of Mexico moisture transport along the Great Plains low-level jet, which has not strengthened uniformly over the historical record (Bukovsky & Mearns, 2020), and warm-season rainfall has decreased in line with the regional drying signal noted across the central Canadian Prairies in our trend analysis. The convergence of three methodologically independent attribution diagnostics with the precipitation, temperature, and streamflow trends confirms that the central-south non-detection is not produced by the bootstrap procedure or by the choice of attribution framework. Rather, it reflects a sub-regional precipitation regime in which the historical record contains no detectable strengthening of 29-day extremes under warming. Reading the sub-region station by station, the same two stations flagged in Sections 4.4.1 and 4.4.2 yield 90% confidence intervals entirely below unity in both observational methods, which we interpret as significant natural-variability detections; at

49.56°N, 99.07°W, however, the GCM-based method flags the same location as significantly anthropogenic (RR = 2.35; 90% CI 1.11 to 8.19), a divergence between the observed and modeled signal that reflects the limited statistical power available at the point scale. The remaining six central-south stations show no statistically detectable anthropogenic signal, which we read as a non-detection rather than positive evidence of natural-variability dominance.

4.5.2 Hydroclimatological Conditions for the 2022 Flood

Understanding the hydroclimatological conditions that preceded and triggered the 2022 flood provides essential context for interpreting the attribution results. The factors that contributed to the Spring 2022 flood were above-average winter snow accumulation (Hydrologic Forecast Centre & Emergency Measures Organization, 2022) and an intense extratropical cyclone on 22 to 24 April 2022, the most intense of the spring sequence. Downstream blocking or ridging may have slowed the system's movement. The event emerged from the interaction between antecedent snowpack conditions during October 2021 to March 2022 and meteorological forcing during March to May 2022. Daily observations from Indian Bay were evaluated relative to the 1990 to 2020 climatological baseline to characterize these drivers.

During the winter season, Indian Bay recorded substantially above-average snow accumulation, indicating enhanced solid-phase water storage prior to the onset of melt, with a mean daily snow depth of approximately 24 cm and a maximum daily snow depth of 61 cm, the latter corresponding to the 90th percentile of the historical November to March seasonal-maximum distribution. Trigger conditions during March to May 2022 further amplified flood potential. Mean temperatures during this period were cooler than the climatological baseline at Indian Bay, falling near the 18th percentile in early spring. The winter snowpack melted through March under these conditions, leaving the soils wet and with reduced infiltration capacity entering April. A fresh

snowfall on 14 April was subsequently melted by the intense 22 to 24 April storm, so that rain-on-snow fell onto already-saturated ground, leading to rapid runoff. Rainfall during March to May was exceptionally high at Indian Bay, with approximately 212 mm corresponding to the 97th percentile, placing spring 2022 among the wetter spring seasons in the historical record at this station.

The synoptic conditions over winter 2021 to 2022 coincided with a moderate La Niña phase of the El Niño-Southern Oscillation, which over the Canadian Prairies is generally associated with cooler winter and spring temperatures and modest positive precipitation anomalies (B. Bonsal et al., 2001). Large-scale climate variability may also influence winter snowpack in western Canada, although its role in Prairie Spring flooding is less well quantified than for British Columbia. The meteorological setup that produced the 2022 event is therefore broadly consistent with established cold-phase teleconnection effects on Prairie hydroclimate, although quantifying the relative contributions of ENSO and the Pacific Decadal Oscillation to the event probability is beyond the scope of this study. Combined, the hydroclimatological evidence indicates that the 2022 flood was not driven by a single meteorological extreme but by the compound interaction of above-normal antecedent snow storage, cool early-spring conditions, and exceptional late-April rain-on-snow.

Blöschl et al. (2019) showed that climate change can both increase and decrease flood hazard depending on the dominant flood-generation mechanism and its regional sensitivity to warming. In our Canadian Prairies and Shield study area, the combined effects of increasing precipitation and low spring temperatures elevated the risk of flooding in the region. The Spring 2022 event was characterized by an exceptionally rare, multi-week precipitation accumulation, yet the resulting flood magnitude at the Whitemouth River was of more moderate historical rarity. This mismatch reflects the nonlinear hydrological controls typical of cold-region basins, where runoff

generation depends critically on antecedent snowpack, soil-moisture state, and frozen-ground conditions. Three characteristics of the Whitemouth sub-basin explain the damping. First, the 189.4 mm accumulation is a point value at Indian Bay, while the flood responds to rainfall averaged over the approximately 3,750 km² drainage area, and the basin-average input was less rare than the station maximum. Second, the sub-basin has low-gradient terrain with forest, wetland, and depressional storage, which holds part of the input and releases it slowly (Hayashi et al., 2016; He et al., 2023). Third, some of the input was retained in the snowpack, and infiltration increased as frozen ground thawed during the event, so the runoff was spread over the 27-day window instead of producing a sharper peak (Fang & Pomeroy, 2007). Saturated soils and rain-on-snow conditions promoted efficient runoff generation, yet these storage and routing processes limited peak-flow amplification. This means that even if anthropogenic warming has raised the likelihood of prolonged heavy precipitation, the effect does not necessarily translate into a proportionate change in flood magnitude, a finding consistent with previous studies emphasizing that flood attribution cannot be inferred directly from precipitation attribution, particularly in snow-dominated or seasonally frozen basins (Scussolini et al., 2024).

4.5.3 Broader Implications and Future Directions

From a risk-management perspective, the findings suggest that anthropogenic climate change may be increasing the background likelihood of hydrometeorological conditions conducive to spring flooding, even if the magnitude of individual floods remains difficult to attribute with confidence. This distinction is important for adaptation planning: an increased frequency of prolonged wet periods or rain-on-snow conditions can elevate flood risk without necessarily producing record-breaking peak flows.

Future progress in cold-region flood attribution will require several advances. These include higher-resolution climate simulations capable of resolving mesoscale precipitation processes, improved representation of snow and frozen-soil dynamics in hydrological models, and longer, more spatially dense observational records. Expanding attribution analyses across multiple basins and events, rather than focusing on single extremes, will also help improve statistical power and clarify the conditions under which anthropogenic signals emerge in flood behavior. These results highlight the need for integrated attribution frameworks that jointly resolve atmospheric forcing and cold-region hydrological response.

4.6 Data Availability

No new observational data were generated for this study; all input datasets are publicly available and cited below. The CMIP6-forced streamflow simulations produced for the attribution analysis were generated by applying the LSTM and Raven models to the listed CMIP6 forcings and can be reproduced using the publicly available model frameworks cited below. Precipitation and temperature data are from the Serially Complete Dataset for North America (SCDNA; Tang et al., 2020). Global mean land temperature anomalies used in the scaling method were obtained from the NOAA Global Surface Temperature dataset (NOAAGlobalTemp; NOAA National Centers for Environmental Information, 2025). Additional station observations are from the Environment and Climate Change Canada Historical Climate Data archive (Environment and Climate Change Canada, 2024). Streamflow records are from the Water Survey of Canada National Water Data Archive, HYDAT (Water Survey of Canada, 2024). Snow depth records are from the Canadian Historical Daily Snow Depth Database (Environment and Climate Change Canada, 2025). CMIP6 historical and hist-nat (DAMIP) simulations from CNRM-CM6-1, BCC-CSM2-MR, IPSL-CM6A-LR, and NorESM2-LM are available through the Earth System Grid Federation

(World Climate Research Programme, 2025); these datasets can be located on any active ESGF node by searching the source model, experiment (historical and hist-nat), and variable facets. The Raven hydrological modeling framework (Craig et al., 2020) and the NeuralHydrology LSTM library (Kratzert et al., 2022) are openly available.

Chapter 5 Conclusions and Future Directions

5.1 General Summary

Based on a multi-method probabilistic event attribution of the Spring 2022 flood, this thesis quantifies how anthropogenic climate change is expressed in the precipitation, streamflow, and compound hydroclimatic conditions of cold-region Prairie and Shield sub-basins. In the analysis three attribution methods (scaling, time-slice, and GCM-based) were used considering 19 precipitation stations and 28 streamflow gauges, streamflow simulation at the Whitemouth sub-basin with a data-driven LSTM and the process-based Raven model forced with four CMIP6 DAMIP models, and a copula-based assessment of the joint snow-temperature preconditioning at Indian Bay. The key findings, summarized below in the order of the objectives of Section 1.2, can serve as guidance for future flood attribution studies in similar cold-region settings.

The event itself proved to be far rarer in the atmosphere than in the river. The duration-optimized event definition identified a 29-day precipitation accumulation of 189.4 mm at Indian Bay, the largest value in the 73-year record, with a GEV-estimated return period of roughly 675 years, whereas the corresponding 27-day mean discharge of $141.6 \text{ m}^3 \text{ s}^{-1}$ at the Whitemouth River gauge had a return period of only about 25.7 years. The sub-basin therefore converted an exceptionally rare precipitation input into a moderately rare flood. This damping reflects the storage-controlled runoff response of the region, in which snowpack, frozen soils, and depressional storage absorb part of the meteorological input, and it means that attribution studies that treat precipitation rarity as equivalent to flood rarity will overstate the anthropogenic contribution to flood hazard.

For precipitation, all three attribution methods indicate that anthropogenic forcing has increased the likelihood of 2022-type multi-week spring precipitation. Between 5 and 8 of the 19 stations are significant at the 90% level depending on the method, and every method reproduces the same east-west gradient (Appendix Figure C3), with the strongest and most consistent signal over the Canadian Shield portion of the study area and a weak or absent signal over the Red River Valley. The direction of the signal is therefore robust, although its magnitude is constrained only to the range spanned by the three methods. For streamflow, in contrast, no method yields a statistically significant signal at the Whitemouth River: the risk ratio is 1.00 for scaling, 2.52 for time-slice, and 1.45 (LSTM) and 3.14 (Raven) for the GCM-based method, with all intervals spanning unity. This is consistent across all three attribution methods and both hydrological models.

The compound analysis showed that the 2022 preconditioning was jointly rare rather than individually extreme. April temperature and peak snow depth correspond to return periods of only about 13 and 16 years on their own, but their joint occurrence corresponds to roughly 60 years, because cooler April temperatures and deeper snowpacks tend to occur together. Comparing the periods before and after 1985, the change in the joint probability of 2022-like conditions is not statistically significant; however, the dependence between the two drivers strengthened, with the best-fitting copula shifting from the Frank to the Gaussian family, while the probability of a peak snow depth as high as in 2022 decreased on its own.

Taken together, these results indicate a partial decoupling between precipitation attribution and flood attribution. Warming has likely increased the background probability of prolonged flood-conducive spring precipitation, particularly over eastern Manitoba and northwestern Ontario, while flood magnitude remains governed by antecedent snow storage, melt timing, and sub-basin

storage and carries no detectable anthropogenic signal. A non-detection in streamflow should not be read as evidence of unchanged flood hazard: the likelihood of heavy spring precipitation is increasing while the long-term probability of a deep peak snowpack is decreasing, consistent with documented snow-cover declines across the northern mid-latitudes (Brown & Mote, 2009), and these opposing changes can keep the net signal in flood magnitude near zero. At the same time, the 2022 snowpack itself was well above the climatological mean, and years in which a deep snowpack aligns with a cool spring can still produce major floods.

5.2 Recommendations for Future Research

Although this work provides important insight into the influence of anthropogenic climate change on the precipitation, streamflow, and compound drivers of the Spring 2022 flood, several future research directions can be considered to either strengthen the conclusions of this study or to develop new knowledge related to cold-region flood attribution and flood risk management in Canada.

- The current analysis is focused to the Whitemouth River sub-basin as the focal catchment. Extending the detection and attribution framework to a broader network of watersheds across southern and central Manitoba would strengthen the generalisability of the findings and provide a more spatially comprehensive understanding of how anthropogenic climate change is altering flood likelihood and magnitude across the region.
- The probabilistic event attribution framework employed in this thesis is fundamentally statistical in nature. Future work should complement this approach with dynamical detection and attribution methods, such as fingerprint-based techniques, which provide a physically grounded basis for identifying the anthropogenic signal in observed climate

trends. The inclusion of dynamical modeling approaches would also improve understanding of how storm systems and their associated precipitation patterns are evolving under climate change.

- Future research should extend the compound attribution analysis to explicitly examine the interacting flood-generating mechanisms that are particularly relevant in cold-region Prairie basins, including rain-on-snow events, in which rainfall on an existing snowpack under frozen ground conditions can produce disproportionately large flood responses. Characterising how the frequency and intensity of such compound triggering mechanisms are changing under anthropogenic warming would improve the understanding of evolving flood risk in the region.
- The hydrological models applied in this thesis do not fully account for the fill-and-spill dynamics and depression storage that are characteristic features of Prairie hydrology, where numerous topographic depressions temporarily store and slowly release water, fundamentally altering the relationship between precipitation and runoff. Incorporating these processes explicitly into the hydrological modeling framework would improve the physical realism of streamflow simulations and is expected to reduce uncertainty in the flood attribution results, particularly for events where antecedent storage conditions play a determining role in flood generation.

These advances would strengthen flood risk assessment and support climate-resilient water management in cold-region environments across the Canadian Prairies and beyond.

References

- Abdelmoaty, H. M., Rajulapati, C. R., Nerantzaki, S. D., & Papalexiou, S. M. (2025). Bias-corrected high-resolution temperature and precipitation projections for Canada. *Scientific Data*, 12(1), 191. <https://doi.org/10.1038/s41597-025-04396-z>
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716–723. <https://doi.org/10.1109/TAC.1974.1100705>
- Atashi, V., Barati, R., & Lim, Y. H. (2023). Development of a distributed nonlinear Muskingum model by considering snowmelt effects for flood routing in the Red River. *Scientific Reports*, 13(1), 21356. <https://doi.org/10.1038/s41598-023-48895-8>
- Bevacqua, E., Maraun, D., Hobæk Haff, I., Widmann, M., & Vrac, M. (2017). Multivariate statistical modelling of compound events via pair-copula constructions: analysis of floods in Ravenna (Italy). *Hydrology and Earth System Sciences*, 21(6), 2701–2723. <https://doi.org/10.5194/hess-21-2701-2017>
- Blöschl, G., Hall, J., Viglione, A., Perdigão, R. A. P., Parajka, J., Merz, B., Lun, D., Arheimer, B., Aronica, G. T., Bilibashi, A., Boháč, M., Bonacci, O., Borga, M., Čanjevac, I., Castellarin, A., Chirico, G. B., Claps, P., Frolova, N., Ganora, D., ... Živković, N. (2019). Changing climate both increases and decreases European river floods. *Nature*, 573(7772), 108–111. <https://doi.org/10.1038/s41586-019-1495-6>
- Boluwade, A., & Rasmussen, P. F. (2015). Frequency of floods in a changing climate: a case study from the Red River in Manitoba, Canada. *Proceedings of the International Association of Hydrological Sciences*, 371, 83–88. <https://doi.org/10.5194/piahs-371-83-2015>

- Bonsal, B. R., Peters, D. L., Seglenieks, F., Rivera, A., & Berg, A. (2019). Changes in freshwater availability across Canada. In E. Bush & D. S. Lemmen (Eds.), *Canada's Changing Climate Report* (pp. 261–342). Government of Canada.
<https://changingclimate.ca/CCCR2019/chapter/6-0/>
- Bonsal, B., Shabbar, A., & Higuchi, K. (2001). Impact of low frequency variability modes on Canadian winter temperature. *International Journal of Climatology*, 21, 95–108.
<https://doi.org/10.1002/joc.590>
- Brown, R. D., & Mote, P. W. (2009). The Response of Northern Hemisphere Snow Cover to a Changing Climate. *Journal of Climate*, 22(8), 2124–2145.
<https://doi.org/10.1175/2008JCLI2665.1>
- Bukovsky, M. S., & Mearns, L. O. (2020). Regional climate change projections from NA-CORDEX and their relation to climate sensitivity. *Climatic Change*, 162(2), 645–665.
<https://doi.org/10.1007/s10584-020-02835-x>
- Bush, E., & Lemmen, D. S. (2019). *Canada's Changing Climate Report*. Government of Canada.
<https://changingclimate.ca/CCCR2019/>
- Buttle, J. M., Allen, D. M., Caissie, D., Davison, B., Hayashi, M., Peters, D. L., Pomeroy, J. W., Simonovic, S., St-Hilaire, A., & Whitfield, P. H. (2016). Flood processes in Canada: Regional and special aspects. *Canadian Water Resources Journal / Revue Canadienne Des Ressources Hydriques*, 41(1–2), 7–30. <https://doi.org/10.1080/07011784.2015.1131629>
- Canadian Climate Institute. (2024). *FACT SHEET: Climate change and flooding*.
<https://climateinstitute.ca/news/fact-sheet-climate-change-and-flooding/>

- Cattiaux, J., & Ribes, A. (2018). Defining Single Extreme Weather Events in a Climate Perspective. *Bulletin of the American Meteorological Society*, 99(8), 1557–1568. <https://doi.org/10.1175/BAMS-D-17-0281.1>
- Coles, S. (2001). *An Introduction to Statistical Modeling of Extreme Values*. Springer. <https://doi.org/10.1007/978-1-4471-3675-0>
- Commission for Environmental Cooperation. (2010). Land cover, 2010 (MODIS, 250 m) [Dataset]. North American Land Change Monitoring System. <http://www.cec.org/north-american-environmental-atlas/land-cover-2010-modis-250m/>
- Craig, J. R., Brown, G., Chlumsky, R., Jenkinson, R. W., Jost, G., Lee, K., Mai, J., Serrer, M., Sgro, N., Shafii, M., Snowdon, A. P., & Tolson, B. A. (2020). Flexible watershed simulation with the Raven hydrological modelling framework. *Environmental Modelling & Software*, 129, 104728. <https://doi.org/10.1016/j.envsoft.2020.104728>
- Dhasmana, M., Mondal, A., & Zachariah, M. (2023). On the role of climate change in the 2018 flooding event in Kerala. *Environmental Research Letters*, 18. <https://doi.org/10.1088/1748-9326/ace6c0>
- Dong, S., Sun, Y., Li, C., Zhang, X., Min, S.-K., & Kim, Y.-H. (2021). Attribution of Extreme Precipitation with Updated Observations and CMIP6 Simulations. *Journal of Climate*, 34(3), 871–881. <https://doi.org/10.1175/JCLI-D-19-1017.1>
- Doocy, S., Daniels, A., Murray, S., & Kirsch, T. (2013). The Human Impact of Floods: A Historical Review of Events 1980-2009 and Systematic Literature Review. *PLoS Currents*, 5. <https://doi.org/10.1371/currents.dis.f4deb457904936b07c09daa98ee8171a>
- Dumanski, S., Pomeroy, J., & Westbrook, C. (2015). Hydrological regime changes in a Canadian Prairie basin. *Hydrological Processes*, 29, 3893–3904. <https://doi.org/10.1002/hyp.10567>

- Efron, B. (1979). Bootstrap Methods: Another Look at the Jackknife. *Annals of Statistics*, 7, 1–26.
<https://doi.org/10.1214/aos/1176344552>
- Environment and Climate Change Canada. (2023). *Canada's top 10 weather stories of 2022*. Government of Canada. <https://www.canada.ca/en/environment-climate-change/services/top-ten-weather-stories/2022.html>
- Environment and Climate Change Canada. (2024). Historical climate data [Dataset]. Government of Canada. https://climate.weather.gc.ca/historical_data/search_historic_data_e.html
- Environment and Climate Change Canada. (2025). Canadian historical daily snow depth database [Dataset]. Government of Canada Open Data Portal. <https://data-donnees.az.ec.gc.ca/data/climate/scientificknowledge/canadian-historical-daily-snow-depth-database>
- Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, 9(5), 1937–1958.
<https://doi.org/10.5194/gmd-9-1937-2016>
- Fang, X., & Pomeroy, J. (2007). Snowmelt runoff sensitivity analysis to drought on the Canadian Prairies. *Hydrological Processes*, 21, 2594–2609. <https://doi.org/10.1002/hyp.6796>
- Gauch, M., Kratzert, F., Klotz, D., Nearing, G., Lin, J., & Hochreiter, S. (2021). Rainfall–runoff prediction at multiple timescales with a single Long Short-Term Memory network. *Hydrology and Earth System Sciences*, 25(4), 2045–2062. <https://doi.org/10.5194/hess-25-2045-2021>
- Gillett, N. P., Cannon, A. J., Malinina, E., Schnorbus, M., Anslow, F., Sun, Q., Kirchmeier-Young, M., Zwiers, F., Seiler, C., Zhang, X., Flato, G., Wan, H., Li, G., & Castellán, A. (2022).

- Human influence on the 2021 British Columbia floods. *Weather and Climate Extremes*, 36, 100441. <https://doi.org/10.1016/j.wace.2022.100441>
- Gillett, N. P., Shiogama, H., Funke, B., Hegerl, G., Knutti, R., Matthes, K., Santer, B. D., Stone, D., & Tebaldi, C. (2016). The Detection and Attribution Model Intercomparison Project (DAMIP~v1.0) contribution to CMIP6. *Geoscientific Model Development*, 9(10), 3685–3697. <https://doi.org/10.5194/gmd-9-3685-2016>
- Government of Manitoba. (2021). *Growing Outcomes in Watersheds (GROW) Guide: A home-grown ecological goods and services program for Manitoba*. https://mbhabitat.ca/wp-content/uploads/2021/09/FINAL_GROW_2021_Guidebook_Sept-2021.pdf
- Gray, D. M., & Male, D. H. (1982). Handbook of Snow: Principles, Processes, Management & Use, edited by D.M. Gray and D.H. Male. *ARCTIC*, 35(2), 346. <https://doi.org/10.14430/arctic2483>
- Gray, D., Toth, B., Zhao, L., Pomeroy, J., & Granger, R. (2001). Estimating Areal Snowmelt Infiltration Into Frozen Soils. *Hydrological Processes*, 15, 3095–3111. <https://doi.org/10.1002/hyp.320>
- Gupta, H. V, Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, 377(1), 80–91. <https://doi.org/10.1016/j.jhydrol.2009.08.003>
- Hayashi, M., van der Kamp, G., & Rosenberry, D. O. (2016). Hydrology of Prairie Wetlands: Understanding the Integrated Surface-Water and Groundwater Processes. *Wetlands*, 36(2), 237–254. <https://doi.org/10.1007/s13157-016-0797-9>

- He, Z., Shook, K., Spence, C., Pomeroy, J. W., & Whitfield, C. (2023). Modelling the regional sensitivity of snowmelt, soil moisture, and streamflow generation to climate over the Canadian Prairies using a basin classification approach. *Hydrology and Earth System Sciences*, 27(19), 3525–3546. <https://doi.org/10.5194/hess-27-3525-2023>
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., & Thépaut, J.-N. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146. <https://doi.org/10.1002/qj.3803>
- Hydrologic Forecast Centre, & Emergency Measures Organization. (2022). *2022 Summer Conditions Report*. https://www.manitoba.ca/floodinfo/pdf/2022/2022_summer_conditions_report.pdf
- Kingma, D., & Ba, J. (2014). Adam: A Method for Stochastic Optimization. *International Conference on Learning Representations*.
- Kirchmeier-Young, M. C., Li, G., Zhang, X., & Wang, X. L. (2025). Attribution of Changes in Canadian Precipitation. *Atmosphere-Ocean*, 63(5), 367–375. <https://doi.org/10.1080/07055900.2025.2545849>
- Kratzert, F., Gauch, M., Nearing, G., & Klotz, D. (2022). NeuralHydrology — A Python library for Deep Learning research in hydrology. *Journal of Open Source Software*, 7, 4050. <https://doi.org/10.21105/joss.04050>
- Kratzert, F., Klotz, D., Shalev, G., Klambauer, G., Hochreiter, S., & Nearing, G. (2019). Towards learning universal, regional, and local hydrological behaviors via machine learning applied

- to large-sample datasets. *Hydrology and Earth System Sciences*, 23(12), 5089–5110.
<https://doi.org/10.5194/hess-23-5089-2019>
- Kumari, S., Haustein, K., Javid, H., Burton, C., Allen, M., Paltán, H., Dadson, S., & Otto, F. (2019). Return period of extreme rainfall substantially decreases under 1.5°C and 2.0°C warming: a case study for Uttarakhand, India. *Environmental Research Letters*, 14.
<https://doi.org/10.1088/1748-9326/ab0bce>
- Kundzewicz, Z. W., Kanae, S., Seneviratne, S. I., Handmer, J., Nicholls, N., Peduzzi, P., Mechler, R., Bouwer, L. M., Arnell, N., Mach, K., Muir-Wood, R., Brakenridge, G. R., Kron, W., Benito, G., Honda, Y., Takahashi, K., & Sherstyukov, B. (2014). Flood risk and climate change: global and regional perspectives. *Hydrological Sciences Journal*, 59(1), 1–28.
<https://doi.org/10.1080/02626667.2013.857411>
- Last, W., & Teller, J. (2004). Paleolimnology of Lake Manitoba: the Lithostratigraphic Evidence. *Geographie Physique et Quaternaire*, 56, 135–154. <https://doi.org/10.7202/009101ar>
- Leach, N., Li, S., Sparrow, S., Van Oldenborgh, G. J., Lott, F., Weisheimer, A., & Allen, M. (2020). Anthropogenic Influence on the 2018 Summer Warm Spell in Europe: The Impact of Different Spatio-Temporal Scales. *Bulletin of the American Meteorological Society*, 101, S41–S46. <https://doi.org/10.1175/BAMS-D-19-0201.1>
- Lees, T., Reece, S., Kratzert, F., Klotz, D., Gauch, M., De Bruijn, J., Kumar Sahu, R., Greve, P., Slater, L., & Dadson, S. J. (2022). Hydrological concept formation inside long short-term memory (LSTM) networks. *Hydrology and Earth System Sciences*, 26(12), 3079–3101.
<https://doi.org/10.5194/hess-26-3079-2022>

- Li, Z.-B., Heuzé, C., Song, J., Crawford, A., Lai, H.-W., & Chen, D. (2025). *All-season analysis of extratropical and Arctic cyclones over the Northern Hemisphere oceans during 1940-2024*. <https://doi.org/10.22541/essoar.175339114.49828217/v1>
- Lynne, M. S., & Dai, A. (2025). Climatology, trends, and variability of cold-season northern hemisphere extratropical cyclone characteristics during 1950–2023 in ERA5. *Climate Dynamics*, 63(12), 464. <https://doi.org/10.1007/s00382-025-07956-6>
- Manitoba Agriculture. (2016). *Subsurface drainage as a water management strategy: adaptive, economic, and environmental considerations*. <https://www.gov.mb.ca/agriculture/environment/soil-management/pubs/subsurface-drainage-as-a-water-management-strategy.pdf>
- Manitoba Infrastructure. (2023). *Red River Floodway Operation Report – Spring 2023*. https://www.gov.mb.ca/mti/wms/wm/pdf/reports/red_river_floodway_operation_report_spring_2023.pdf
- Manitoba Transportation and Infrastructure. (2023, March 24). *Manitoba government continues restoration of infrastructure damaged in spring 2022 flooding*. Province of Manitoba. <https://news.gov.mb.ca/news/index.html?item=58522>
- Martel, J.-L., Arsenault, R., Turcotte, R., Castañeda-Gonzalez, M., Brissette, F., Armstrong, W., Mailhot, E., Pelletier-Dumont, J., Lachance-Cloutier, S., Rondeau-Genesse, G., & Caron, L.-P. (2025). Exploring the ability of LSTM-based hydrological models to simulate streamflow time series for flood frequency analysis. *Hydrology and Earth System Sciences*, 29(19), 4951–4968. <https://doi.org/10.5194/hess-29-4951-2025>
- Merz, B., Blöschl, G., Vorogushyn, S., Dottori, F., Aerts, J. C. J. H., Bates, P. D., Bertola, M., Kemter, M., Kreibich, H., Lall, U., & Macdonald, E. (2021). Causes, impacts and patterns of

- disastrous river floods. *Nature Reviews Earth & Environment*, 2, 592–609.
<https://doi.org/10.1038/s43017-021-00195-3>
- Mudryk, L. R., Derksen, C., Kushner, P. J., & Brown, R. (2015). Characterization of Northern Hemisphere Snow Water Equivalent Datasets, 1981–2010. *Journal of Climate*, 28(20), 8037–8051. <https://doi.org/10.1175/JCLI-D-15-0229.1>
- Musselman, K. N., Lehner, F., Ikeda, K., Clark, M., Prein, A., Liu, C., Barlage, M., & Rasmussen, R. (2018). Projected increases and shifts in rain-on-snow flood risk over western North America. *Nature Climate Change*, 8. <https://doi.org/10.1038/s41558-018-0236-4>
- Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I — A discussion of principles. *Journal of Hydrology*, 10(3), 282–290.
[https://doi.org/10.1016/0022-1694\(70\)90255-6](https://doi.org/10.1016/0022-1694(70)90255-6)
- National Academies of Sciences, Engineering, and Medicine. (2016). *Attribution of Extreme Weather Events in the Context of Climate Change*. The National Academies Press.
<https://doi.org/10.17226/21852>
- NOAA National Centers for Environmental Information. (2025). Global surface temperature dataset (NOAAGlobalTemp) [Dataset]. NOAA NCEI.
<https://www.ncei.noaa.gov/products/land-based-station/noaa-global-temp>
- Otto, F. E. L. (2023). Attribution of Extreme Events to Climate Change. *Annual Review of Environment and Resources*, 48(Volume 48, 2023), 813–828.
<https://doi.org/10.1146/annurev-environ-112621-083538>
- Otto, F., Wiel, K., Van Oldenborgh, G. J., Philip, S., Kew, S., Uhe, P., & Cullen, H. (2018). Climate change increases the probability of heavy rains in Northern England/Southern Scotland like

- those of storm Desmond - A real-time event attribution revisited. *Environmental Research Letters*, 13. <https://doi.org/10.1088/1748-9326/aa9663>
- Pettitt, A. N. (1979). A Non-Parametric Approach to the Change-Point Problem. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 28(2), 126–135. <https://doi.org/10.2307/2346729>
- Philip, S., Kew, S., van Oldenborgh, G. J., Otto, F., Vautard, R., van der Wiel, K., King, A., Lott, F., Arrighi, J., Singh, R., & van Aalst, M. (2020). A protocol for probabilistic extreme event attribution analyses. *Advances in Statistical Climatology, Meteorology and Oceanography*, 6(2), 177–203. <https://doi.org/10.5194/asmo-6-177-2020>
- Ren, L., Zhou, T., & Zhang, W. (2020). Attribution of the record-breaking heat event over Northeast Asia in summer 2018: the role of circulation. *Environmental Research Letters*, 15. <https://doi.org/10.1088/1748-9326/ab8032>
- Salvadori, G., de michele, C., Kottegoda, N., & Rosso, R. (2007). *Extremes in Nature: An Approach Using Copulas*. <https://doi.org/10.1007/1-4020-4415-1>
- Scussolini, P., Luu, L. N., Philip, S., Berghuijs, W. R., Eilander, D., Aerts, J. C. J. H., Kew, S. F., van Oldenborgh, G. J., Toonen, W. H. J., Volkholz, J., & Coumou, D. (2024). Challenges in the attribution of river flood events. *WIREs Climate Change*, 15(3), e874. <https://doi.org/10.1002/wcc.874>
- Shiogama, H., Watanabe, M., Kim, H., & Hirota, N. (2022). Emergent constraints on future precipitation changes. *Nature*, 602(7898), 612–616. <https://doi.org/10.1038/s41586-021-04310-8>

- Sinha, D., Rajulapati, C. R., & Crawford, A. (2025). Concurrent Heatwaves and Droughts in Canada: Spatio-Temporal Changes, Climate Drivers, and Persistence Properties. *Earth's Future*, 13(10), e2025EF006104. <https://doi.org/10.1029/2025EF006104>
- Srivastava, A., Grotjahn, R., & Ullrich, P. (2020). Evaluation of historical CMIP6 model simulations of extreme precipitation over contiguous US regions. *Weather and Climate Extremes*, 29, 100268. <https://doi.org/10.1016/j.wace.2020.100268>
- Stadnyk, T., & Déry, S. (2021). Canadian Continental-Scale Hydrology under a Changing Climate: A Review. *Water*, 13, 906. <https://doi.org/10.3390/w13070906>
- Stephens, M. A. (1974). EDF Statistics for Goodness of Fit and Some Comparisons. *Journal of the American Statistical Association*, 69, 730–737. <https://doi.org/10.1080/01621459.1974.10480196>
- Stott, P. A., Christidis, N., Otto, F. E. L., Sun, Y., Vanderlinden, J.-P., van Oldenborgh, G. J., Vautard, R., von Storch, H., Walton, P., Yiou, P., & Zwiers, F. W. (2016). Attribution of extreme weather and climate-related events. *WIREs Climate Change*, 7(1), 23–41. <https://doi.org/10.1002/wcc.380>
- Tang, G., Clark, M. P., Newman, A. J., Wood, A. W., Papalexiou, S. M., Vionnet, V., & Whitfield, P. H. (2020). SCDNA: a serially complete precipitation and temperature dataset for North America from 1979 to 2018. *Earth System Science Data*, 12(4), 2381–2409. <https://doi.org/10.5194/essd-12-2381-2020>
- Tarek, M., Brissette, F., & Arsenault, R. (2020). Evaluation of the ERA5 reanalysis as a potential reference dataset for hydrological modelling over North America. *Hydrology and Earth System Sciences*, 24, 2527–2544. <https://doi.org/10.5194/hess-24-2527-2020>

- Teegavarapu, R. S. V. (2019). Chapter 1 - Methods for Analysis of Trends and Changes in Hydroclimatological Time-Series. In R. Teegavarapu (Ed.), *Trends and Changes in Hydroclimatic Variables* (pp. 1–89). Elsevier. <https://doi.org/10.1016/B978-0-12-810985-4.00001-3>
- Tellman, B., Sullivan, J. A., Kuhn, C., Kettner, A. J., Doyle, C. S., Brakenridge, G. R., Erickson, T. A., & Slayback, D. A. (2021). Satellite imaging reveals increased proportion of population exposed to floods. *Nature*, *596*(7870), 80–86. <https://doi.org/10.1038/s41586-021-03695-w>
- Teufel, B., Diro, G. T., Whan, K., Milrad, S. M., Jeong, D. I., Ganji, A., Huziy, O., Winger, K., Gyakum, J. R., de Elia, R., Zwiers, F. W., & Sushama, L. (2017). Investigation of the 2013 Alberta flood from weather and climate perspectives. *Climate Dynamics*, *48*(9), 2881–2899. <https://doi.org/10.1007/s00382-016-3239-8>
- Thiessen, A. H. (1911). PRECIPITATION AVERAGES FOR LARGE AREAS. *Monthly Weather Review*, *39*, 1082–1089. [https://doi.org/10.1175/1520-0493\(1911\)39%3C1082b:pafla%3E2.0.co;2](https://doi.org/10.1175/1520-0493(1911)39%3C1082b:pafla%3E2.0.co;2)
- Tolson, B. A., & Shoemaker, C. A. (2007). Dynamically dimensioned search algorithm for computationally efficient watershed model calibration. *Water Resources Research*, *43*(1). <https://doi.org/10.1029/2005WR004723>
- van Oldenborgh, G. J., van der Wiel, K., Kew, S., Philip, S., Otto, F., Vautard, R., King, A., Lott, F., Arrighi, J., Singh, R., & van Aalst, M. (2021). Pathways and pitfalls in extreme event attribution. *Climatic Change*, *166*(1), 13. <https://doi.org/10.1007/s10584-021-03071-7>
- Vincent, L. A., Zhang, X., Brown, R. D., Feng, Y., Mekis, E., Milewska, E. J., Wan, H., & Wang, X. L. (2015). Observed Trends in Canada's Climate and Influence of Low-Frequency

- Variability Modes. *Journal of Climate*, 28(11), 4545–4560. <https://doi.org/10.1175/JCLI-D-14-00697.1>
- Vincent, L. A., Zhang, X., Mekis, É., Wan, H., & Bush, E. J. (2018). Changes in Canada's Climate: Trends in Indices Based on Daily Temperature and Precipitation Data. *Atmosphere-Ocean*, 56(5), 332–349. <https://doi.org/10.1080/07055900.2018.1514579>
- Water Survey of Canada. (2024). National water data archive (HYDAT) [Dataset]. Environment and Climate Change Canada. <https://www.canada.ca/en/environment-climate-change/services/water-overview/quantity/monitoring/survey/data-products-services/national-archive-hydat.html>
- Wehner, M., Duffy, M., Risser, M., Paciorek, C., Stone, D., & Pall, P. (2024). On the uncertainty of long-period return values of extreme daily precipitation. *Frontiers in Climate*, 6. <https://doi.org/10.3389/fclim.2024.1343072>
- Wi, S., & Steinschneider, S. (2022). Assessing the Physical Realism of Deep Learning Hydrologic Model Projections Under Climate Change. *Water Resources Research*, 58. <https://doi.org/10.1029/2022WR032123>
- Wi, S., & Steinschneider, S. (2024). On the need for physical constraints in deep learning rainfall–runoff projections under climate change: a sensitivity analysis to warming and shifts in potential evapotranspiration. *Hydrology and Earth System Sciences*, 28(3), 479–503. <https://doi.org/10.5194/hess-28-479-2024>
- Winter, B., Schneeberger, K., Förster, K., & Vorogushyn, S. (2020). Event generation for probabilistic flood risk modelling: multi-site peak flow dependence model vs. weather-generator-based approach. *Natural Hazards and Earth System Sciences*, 20(6), 1689–1703. <https://doi.org/10.5194/nhess-20-1689-2020>

- World Climate Research Programme. (2025). CMIP6 model output: Historical and DAMIP (historical) simulations [Dataset]. Earth System Grid Federation. <https://esgf-metagrid.cloud.dkrz.de/search?project=CMIP6>
- Xie, H., Li, D., & Xiong, L. (2013). Exploring the ability of the Pettitt method for detecting change point by Monte Carlo simulation. *Stochastic Environmental Research and Risk Assessment*, 28, 1643–1655. <https://doi.org/10.1007/s00477-013-0814-y>
- Zscheischler, J., & Seneviratne, S. (2017). Dependence of drivers affects risks associated with compound events. *Science Advances*, 3, e1700263. <https://doi.org/10.1126/sciadv.1700263>
- Zscheischler, J., Westra, S., Hurk, B., Seneviratne, S., Ward, P., Pitman, A., AghaKouchak, A., Bresch, D., Leonard, M., Wahl, T., & Zhang, X. (2018). Future climate risk from compound events. *Nature Climate Change*, 8. <https://doi.org/10.1038/s41558-018-0156-3>

Appendix A: Supplementary Tables and Figures

Table A1: CMIP6 models used in this study, including nominal horizontal resolution, simulation periods, and modeling centers.

Sr. No.	Model name	Latitude resolution (degree)	Longitude resolution (degree)	Period	Modeling center
1	CNRM- CM6-1	~1.40	~1.40	1950–2022 (Hist + SSP5- 8.5); 1950– 2020 (HistNat)	Centre National de Recherches Météorologiques (CNRM), France
2	BCC- CSM2-MR	~1.12	~1.12	1950–2022 (Hist + SSP5- 8.5); 1950– 2020 (HistNat)	Beijing Climate Center (BCC), China
3	IPSL- CM6A-LR	~1.25	~2.50	1950–2022 (Hist + SSP5- 8.5); 1950– 2020 (HistNat)	Institut Pierre-Simon Laplace (IPSL), France
4	NorESM2- LM	~1.90	~2.50	1950–2022 (Hist + SSP5- 8.5); 1950– 2020 (HistNat)	Norwegian Climate Center (NCC), Norway

Table A2: Schematic mapping of research objectives to the corresponding variables, methodological approaches, and data sources used in this study. The table provides an integrated overview of the analytical framework, linking event characterization, probabilistic attribution, hydrological modeling, and compound hydroclimatic analysis for assessing the role of anthropogenic climate change in a 2022-type spring flood in southern Manitoba.

Objective	Variables	Methods	Data Sources
Event characterization and rarity assessment	Precipitation (RX29D), Streamflow (RX27D), event duration and timing	Extreme Value Theory (GEV), moving-window RXnD analysis, return period estimation, goodness-of-fit tests	Observed precipitation (ECCC, SCDNA), observed streamflow (HYDAT), station-based temperature
Multi-method probabilistic attribution	Precipitation extremes, Streamflow extremes, Risk Ratio (RR), Fraction of Attributable Risk (FAR)	Scaling method (nonstationary GEV with temperature covariate), Time-slice method (change-point detection + stationary GEV), GCM-based attribution (Hist vs HistNat ensembles)	Observations (ECCC, SCDNA, HYDAT), CMIP6 simulations (CNRM-CM6-1, BCC-CSM2-MR, IPSL-CM6A-LR, NorESM2-LM), global temperature records
Hydrological modeling for attribution	Streamflow, basin-scale meteorological forcings (precipitation,	LSTM-based hydrologic modeling (gap-filling and streamflow simulation), model	Observed streamflow (HYDAT), ERA5 reanalysis (basin-averaged precipitation,

(LSTM workflows)	temperature, wind), basin attributes	training/validation/testing, performance metrics (NSE, KGE, RMSE)	temperature, wind), derived basin-scale predictors
Compound hydroclimatic drivers analysis	Snow depth, temperature, precipitation, antecedent wetness, compound indices (snow–rain– temperature interactions)	Joint probability analysis, compound event framework, multivariate/statistical dependence analysis (e.g., copula-based or correlation-based approaches), historical variability assessment	Snow depth datasets (CHDSDD, ECCC), observed precipitation and temperature (ECCC, SCDNA), HYDAT streamflow
Integrated synthesis and implications	Precipitation and streamflow extremes, RR, FAR, compound indicators, uncertainty metrics	Multi-method comparison, uncertainty analysis (bootstrap resampling), cross-method synthesis, physical interpretation of hydroclimatic processes	Combined datasets (observations, CMIP6 simulations, ERA5- based LSTM outputs, derived hydroclimatic indicators)

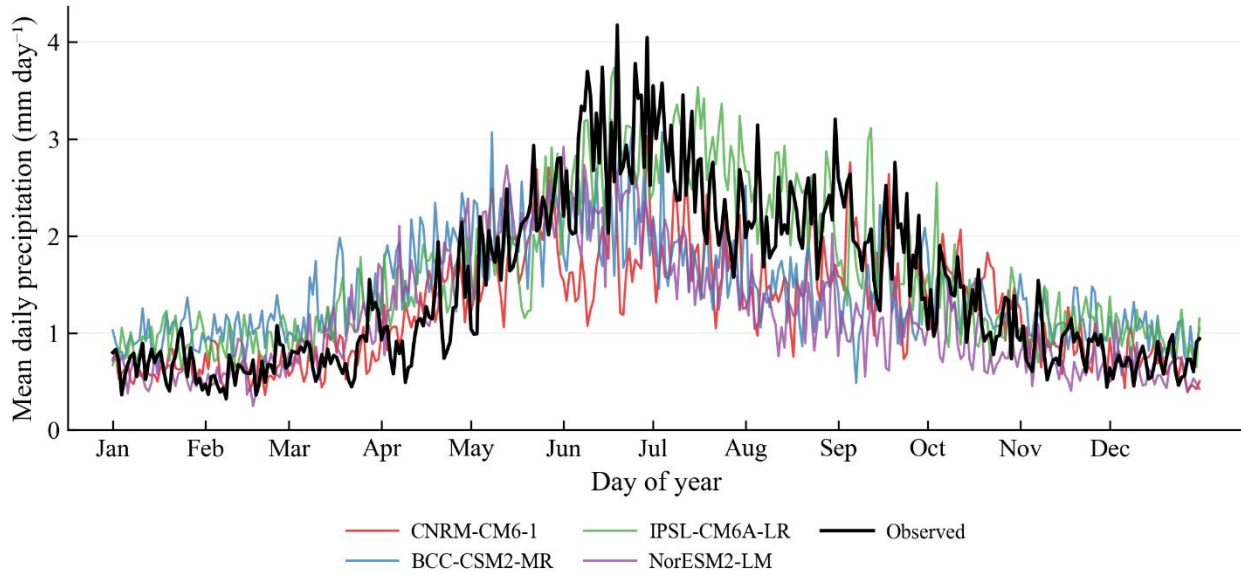


Figure A1: Daily precipitation climatology over the study region for the 1950–2020 period, comparing observed station data (black) with historical simulations from four CMIP6 models: CNRM-CM6-1, BCC-CSM2-MR, IPSL-CM6A-LR, and NorESM2-LM. Each curve represents the multi-station mean daily precipitation for each day of the year, smoothed with a 15-day centered running mean. Observed values are averaged across the 19 precipitation stations; CMIP6 values are averaged across nearest-grid-cell extractions at the corresponding locations.

No bias correction was applied to the CMIP6 output.

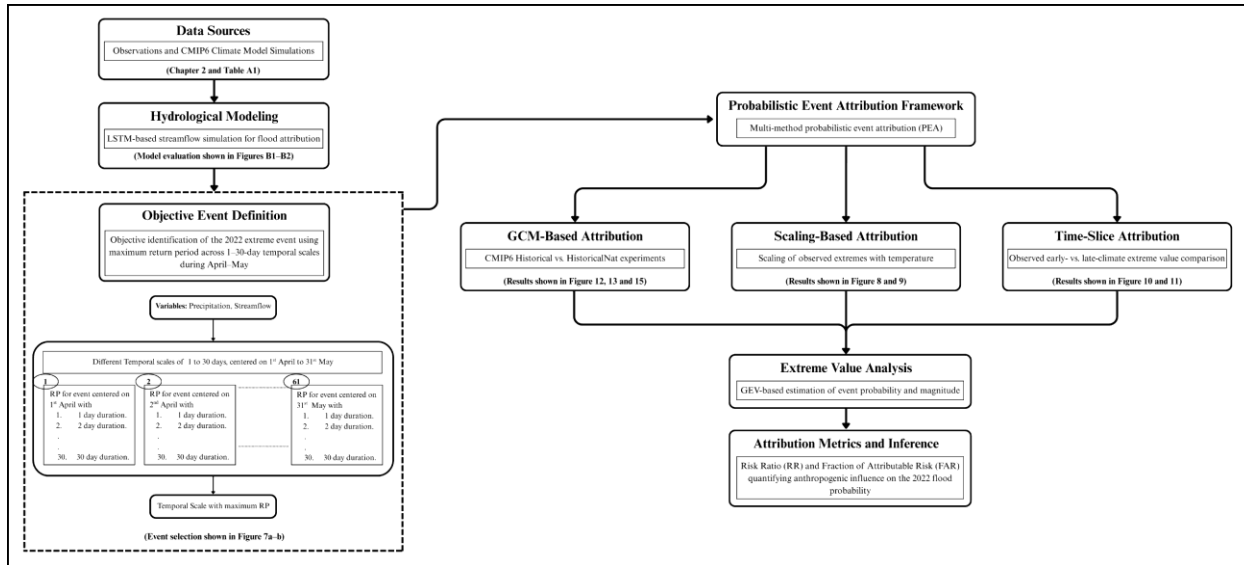


Figure A2: Schematic overview of the probabilistic event attribution framework applied in this study. Following objective event selection, flood attribution is assessed using three approaches: GCM, scaling, and time-slice analysis. Extreme value modeling is used to estimate changes in event probability and associated uncertainty, culminating in a synthesized attribution statement.

Appendix B: LSTM Training and Hyperparameter Diagnostics

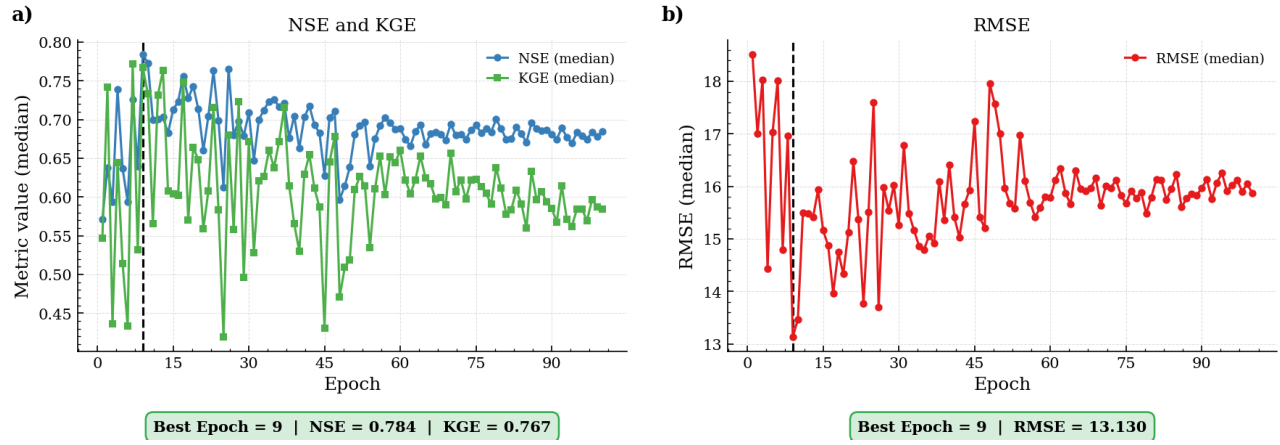


Figure B1: Validation performance across training epochs: (a) median Nash-Sutcliffe Efficiency (NSE) and Kling-Gupta Efficiency (KGE), and (b) median Root Mean Square Error (RMSE).

The dashed vertical line and green box indicate the best epoch selected via multi-objective

ranking.

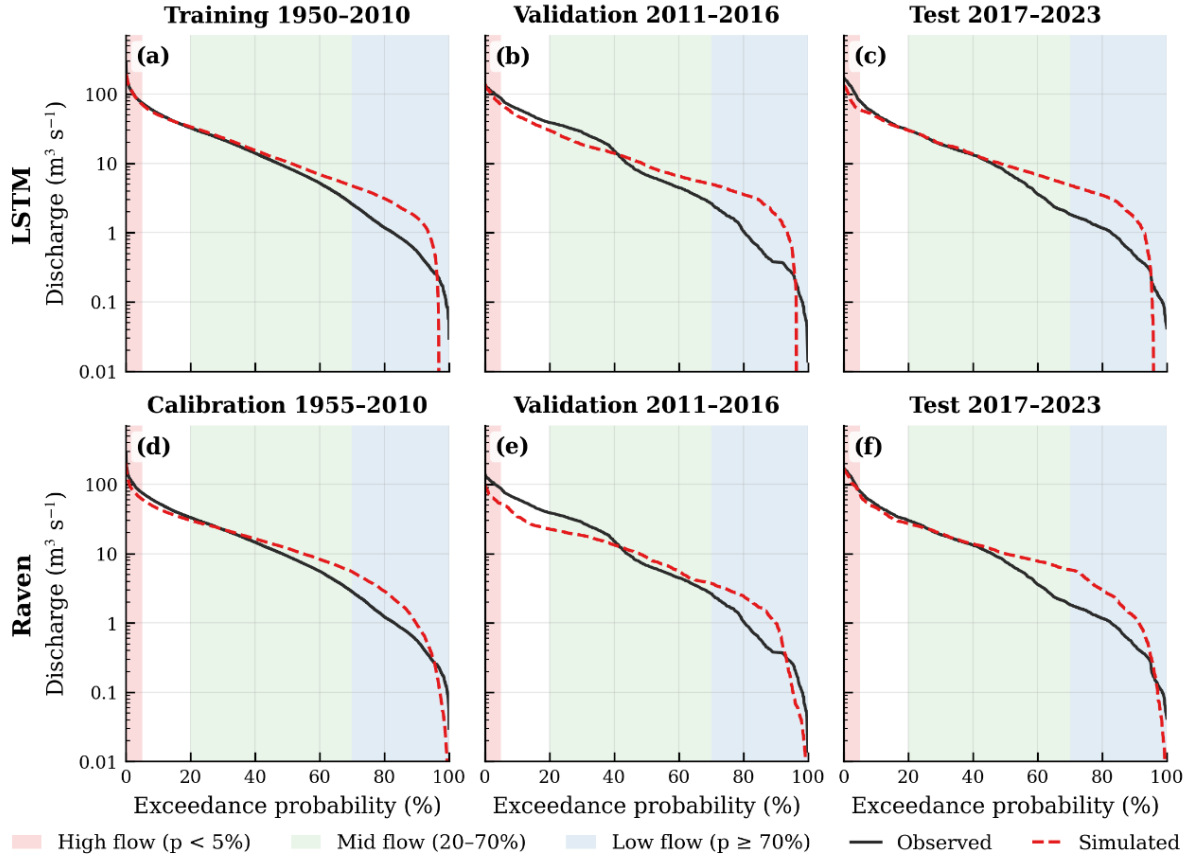


Figure B2: Flow duration curves (FDCs) for station 05PH003 comparing LSTM (a–c) and Raven (d–f) simulated daily streamflow against observations on a logarithmic discharge axis. Panels show (a, d) training/calibration, (b, e) validation, and (c, f) test periods. LSTM periods span 1950–2010, 2011–2016, and 2017–2023; the Raven calibration covers 1955–2010 (1950–1954 treated as warm-up) and its validation run (2011–2023) is subdivided into the validation and test windows for direct comparison with LSTM. Shaded bands denote high-flow ($p < 5\%$), mid-flow (20–70%), and low-flow ($p \geq 70\%$) regimes along the exceedance-probability axis.

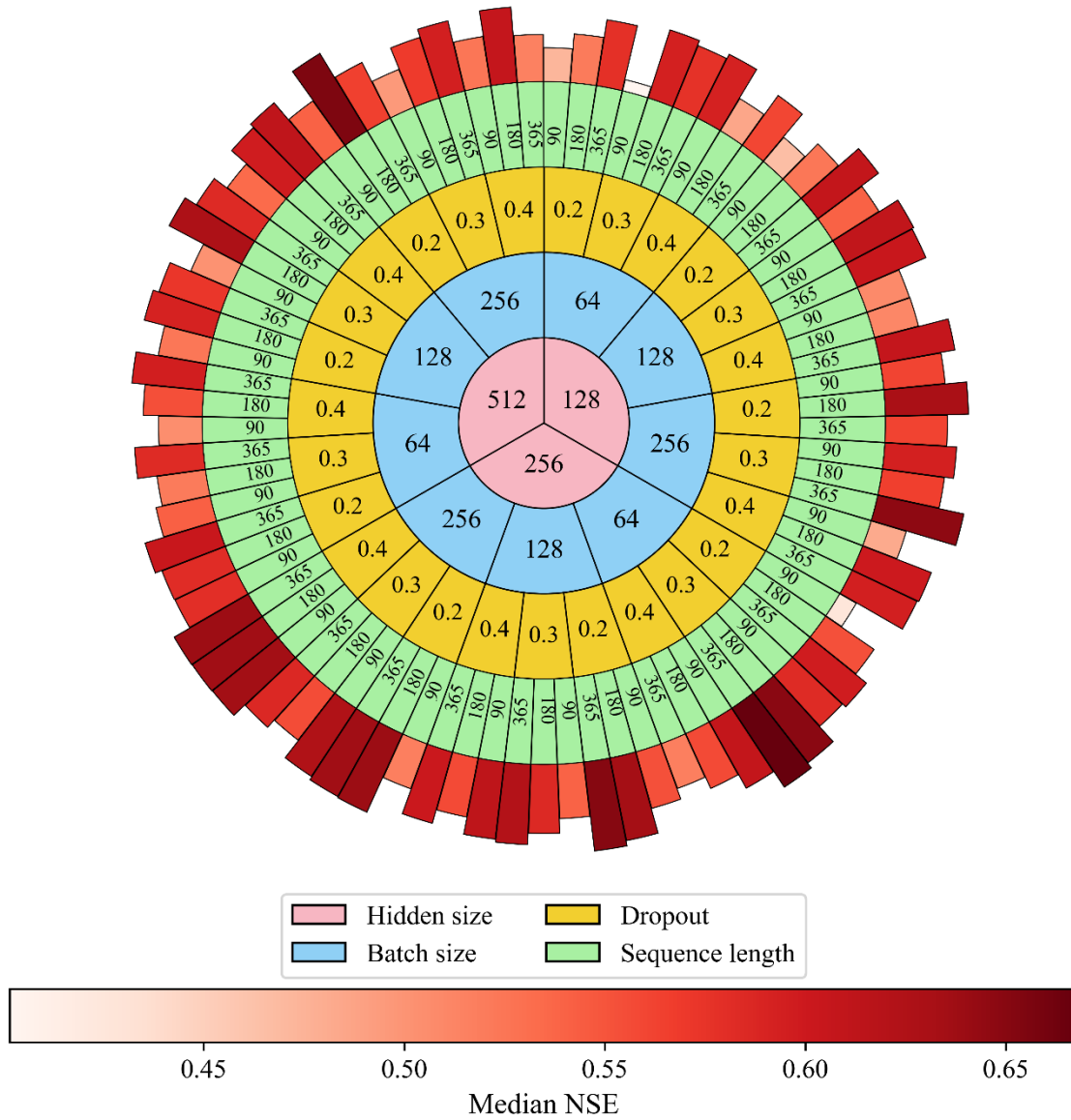


Figure B3: Hyperparameter grid search for the Thiessen-forcing regional LSTM: 81 configurations spanning hidden size, batch size, output dropout, and sequence length. Bar height and color give the best median validation NSE across basins for each configuration. The selected configuration (hidden size 256, batch size 64, dropout 0.3, sequence length 365 days) achieved the highest median validation NSE of the grid (0.67).

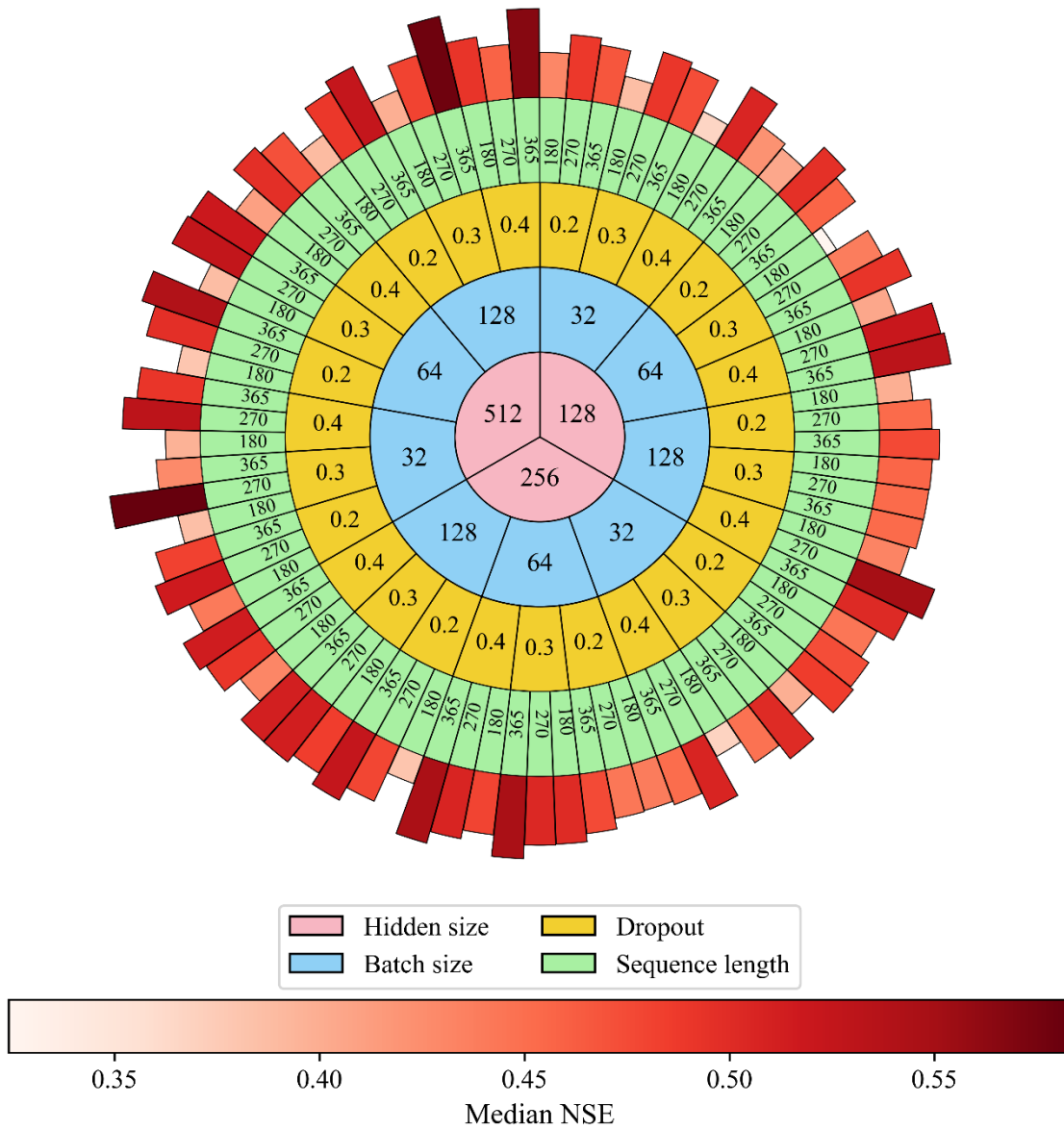


Figure B4: As Figure B3, but for the ERA5-forcing experiment (sequence lengths of 180, 270, and 365 days). The selected configuration (hidden size 512, batch size 128, dropout 0.3, sequence length 365 days; median validation NSE 0.58) was chosen from the top-ranked candidates to retain the 365-day sequence length used in the Thiessen experiment.

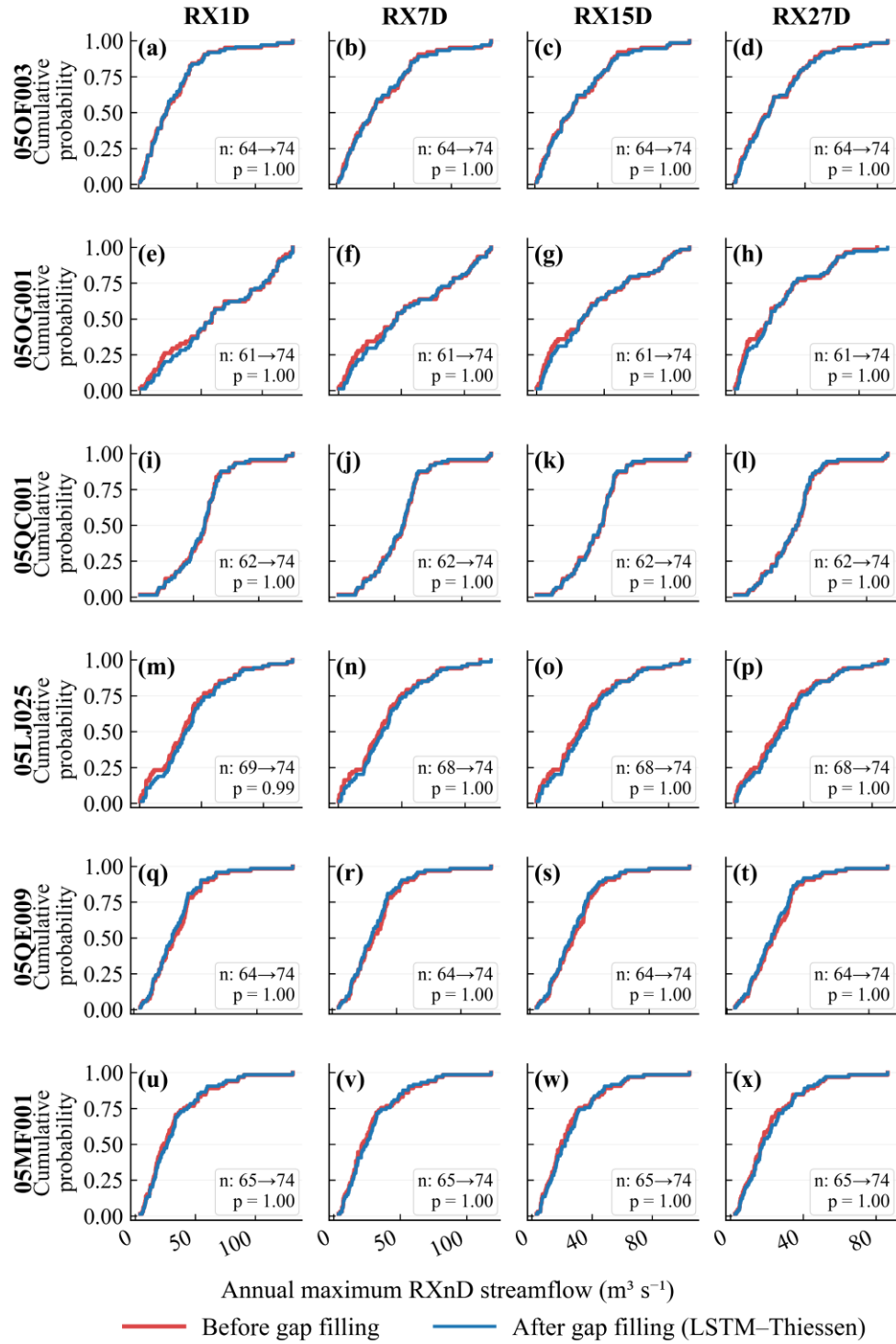


Figure B5: As Figure 6, but for the Thiessen-forced reconstructions; shown for completeness of the forcing comparison

Appendix C: Additional EVT and Attribution Diagnostics

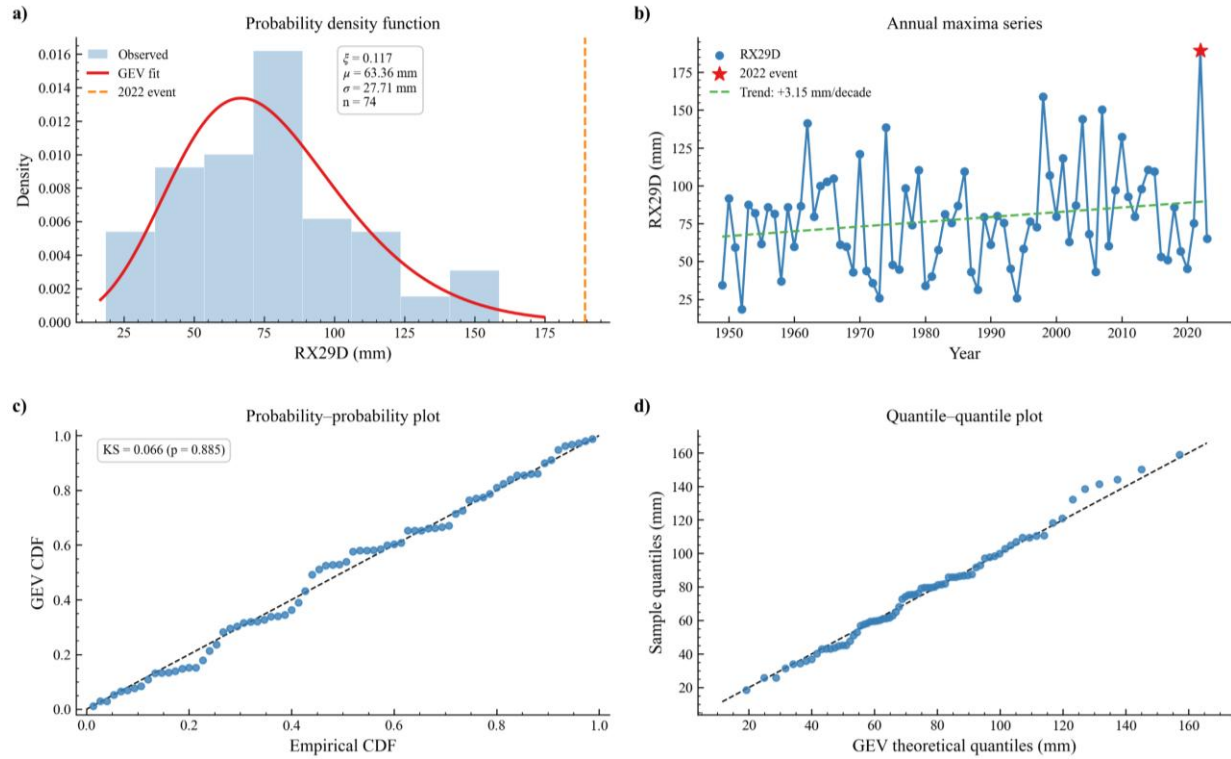


Figure C1: GEV diagnostics for 29-day maximum precipitation (RX29D) at Indian Bay: (a) probability density function with fitted GEV distribution and 2022 event, (b) annual maxima time series with linear trend, (c) probability-probability plot, and (d) quantile-quantile plot.

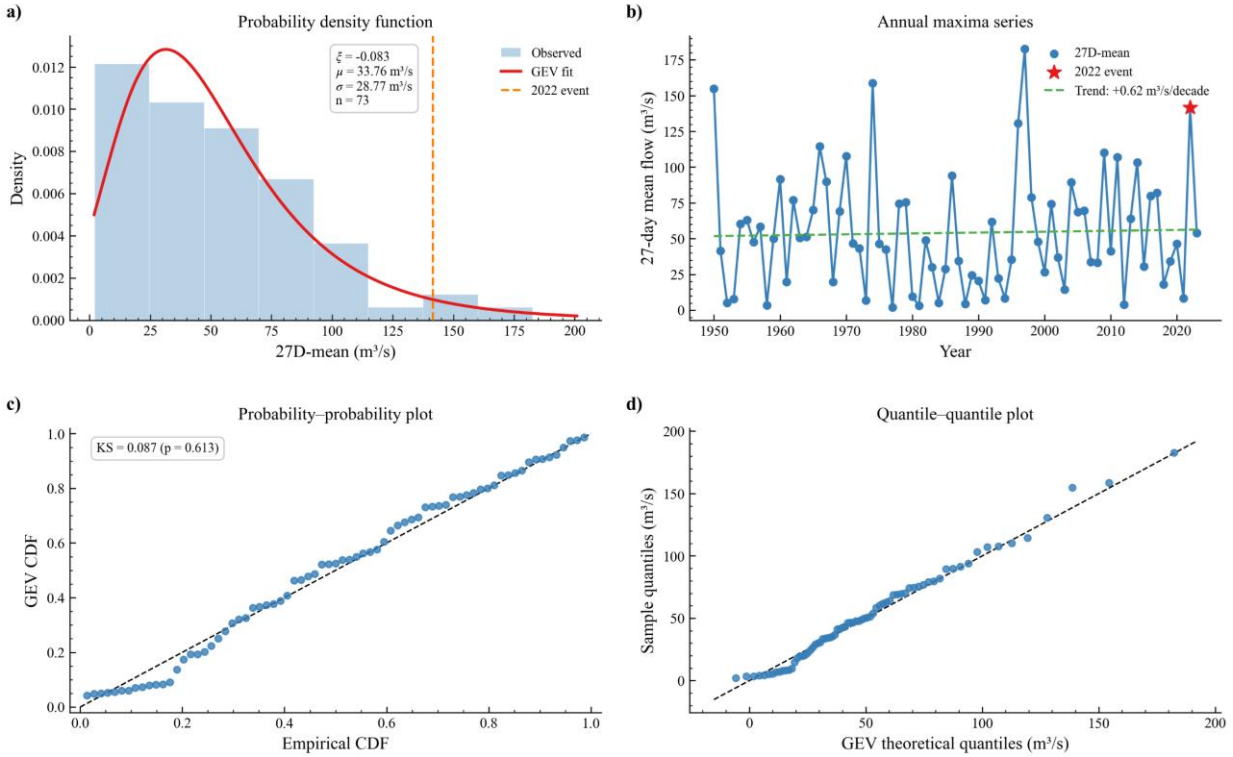


Figure C2: GEV diagnostics for 27-day mean streamflow at the Whitemouth River sub-basin (05PH003): (a) probability density function with fitted GEV distribution and 2022 event, (b) annual maxima time series with linear trend, (c) probability-probability plot, and (d) quantile-quantile plot.

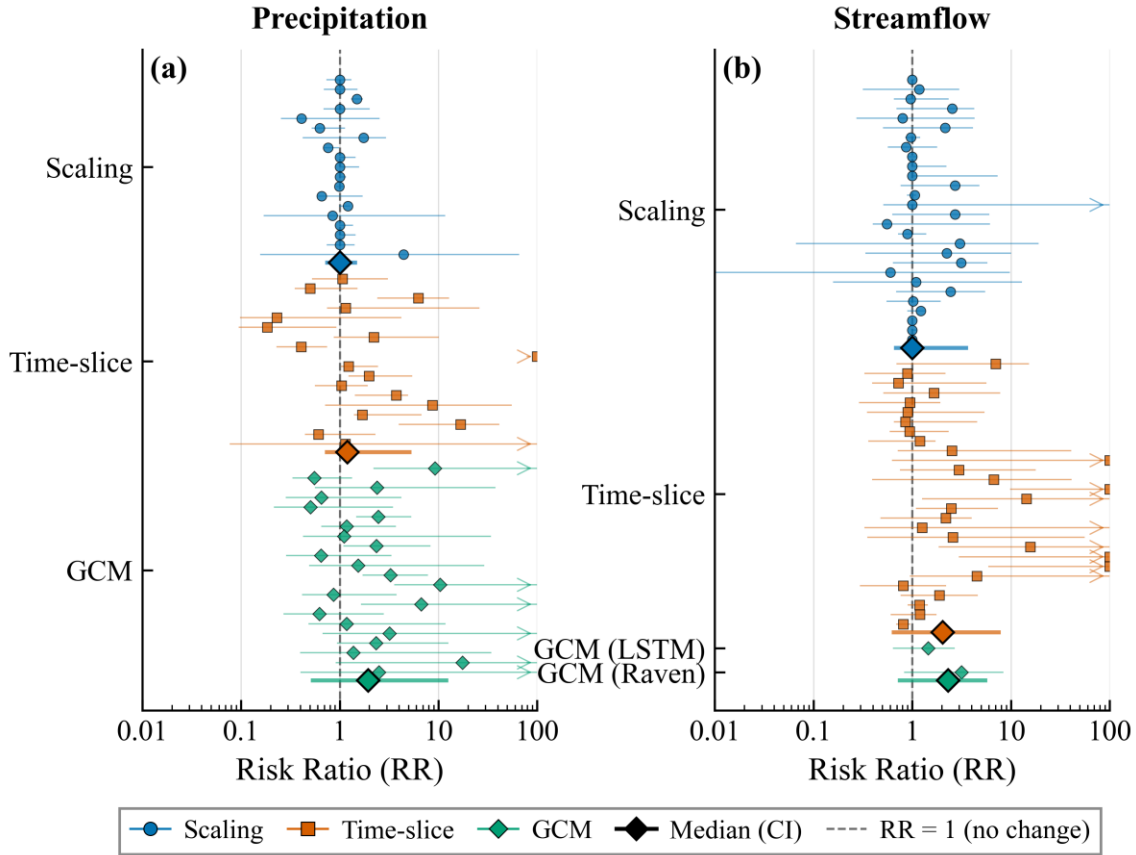


Figure C3: Cross-method comparison of risk ratios (RR) for extreme (a) precipitation (19 stations) and (b) streamflow (28 gauges) events using the Scaling, Time-slice, and GCM-based attribution methods. For streamflow, GCM-based results are shown separately for the LSTM and Raven hydrological models at the Whitemouth River gauge (05PH003), the only gauge at which GCM-forced streamflow was simulated. Colored points represent station-level RR point estimates with their 90% bootstrap confidence intervals; diamonds indicate the across-station median RR with the interquartile range (Q1–Q3) for each method. The dashed vertical line at $RR = 1$ denotes no change in event likelihood. Arrows indicate confidence-interval bounds exceeding the upper x-axis limit.

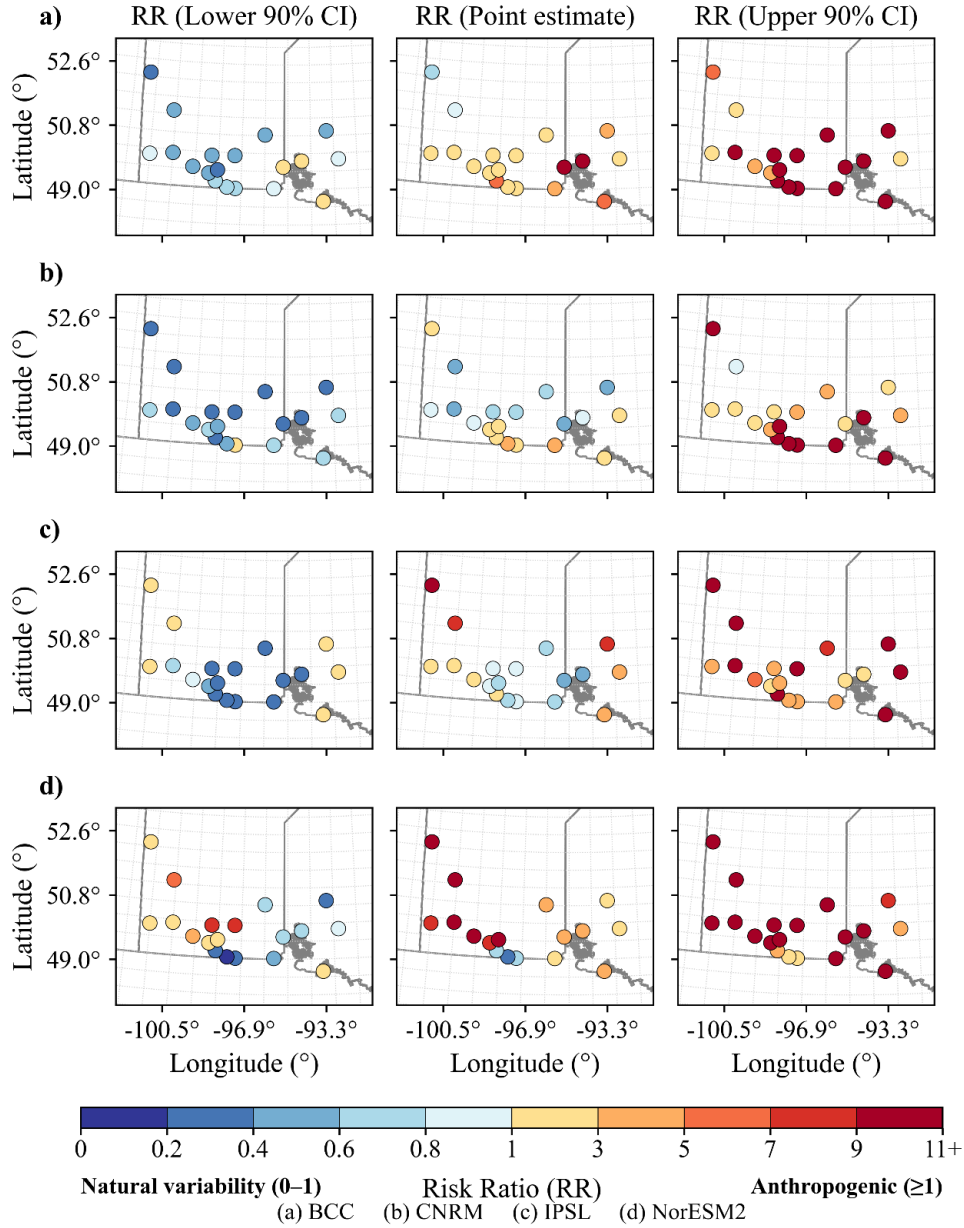


Figure C4: Per-model spatial distributions of station-level risk ratios (RR) derived from the GCM-based attribution method for each of the four CMIP6 models: (a) BCC-CSM2-MR, (b) CNRM-CM6-1, (c) IPSL-CM6A-LR, and (d) NorESM2-LM. For each model, the left, center and right columns display the lower bound of the 90% confidence interval, the point estimate, and the upper bound of the 90% confidence interval, respectively. Colors denote RR magnitude, with blue shades ($RR < 1$) indicating dominance of natural variability and warm shades ($RR \geq 1$) indicating a central estimate of anthropogenic influence.

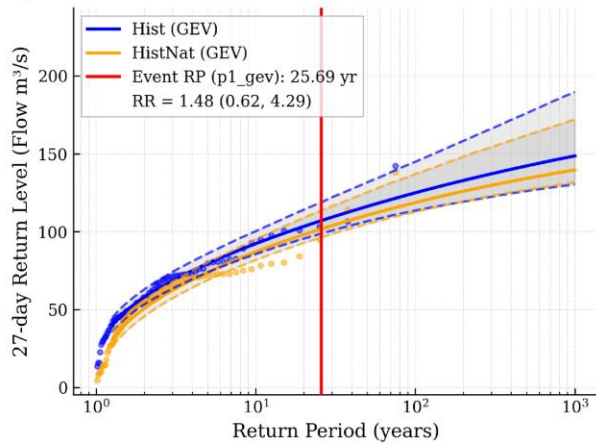
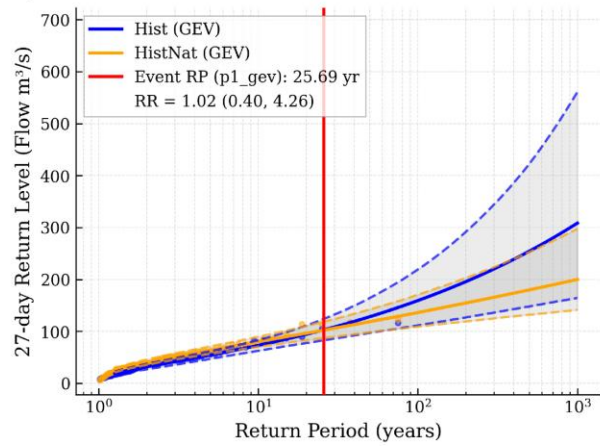
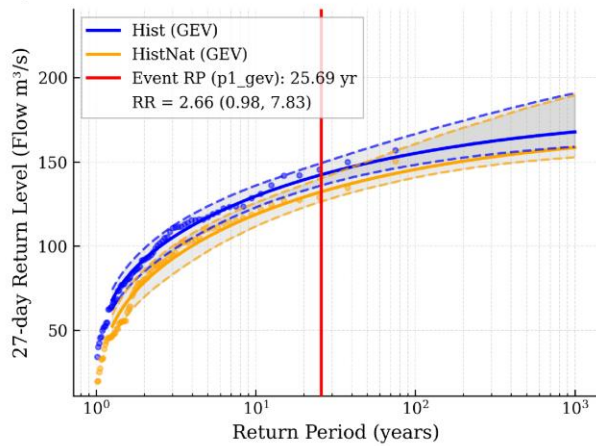
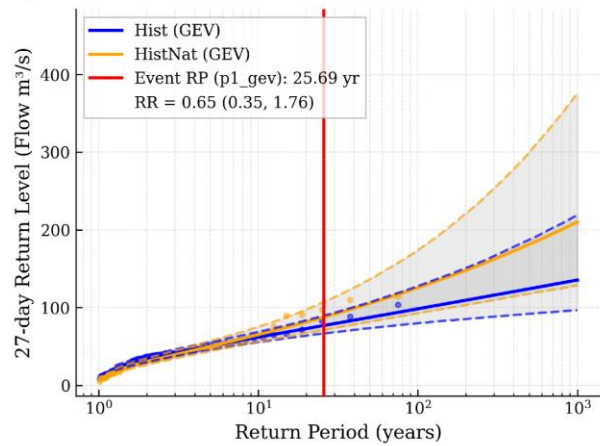
a) BCC-CSM2-MR**b) CNRM-CM6-1****c) IPSL-CM6A-LR****d) NorESM2-LM**

Figure C5: Per-model return-level curves and risk ratios (RR) for annual maximum 27-day mean flow at the Whitemouth River gauge (05PH003), obtained by forcing the LSTM model with each of the four CMIP6 models: (a) BCC-CSM2-MR, (b) CNRM-CM6-1, (c) IPSL-CM6A-LR, and (d) NorESM2-LM. Blue curves represent the historical all-forcing (factual) climate; orange curves represent the historical-natural (counterfactual) climate. Solid lines denote fitted generalized extreme value (GEV) return-level curves, dashed lines indicate the 90% uncertainty bounds, and the vertical red line marks the empirical return period of the 2022 event.

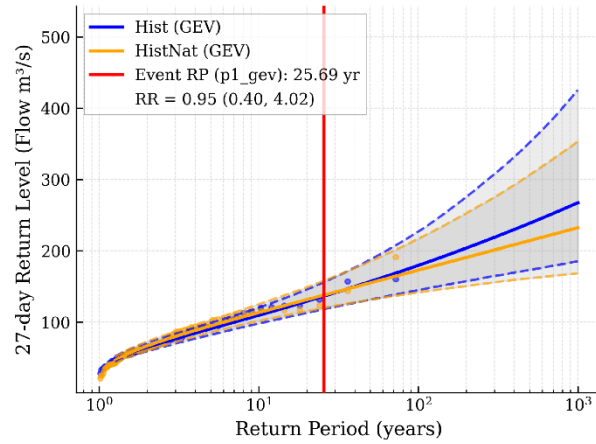
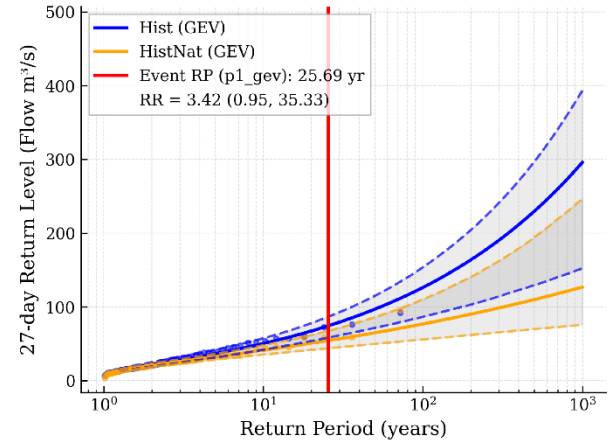
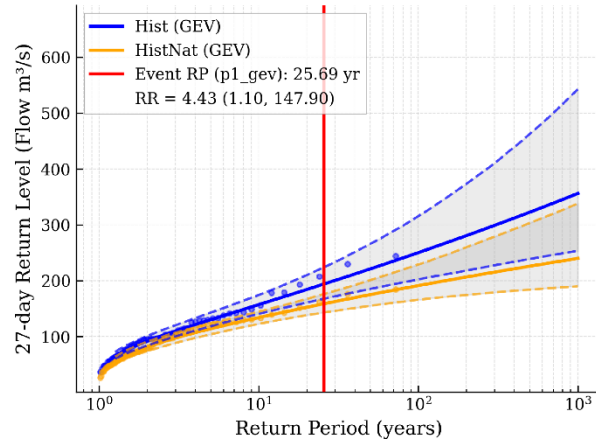
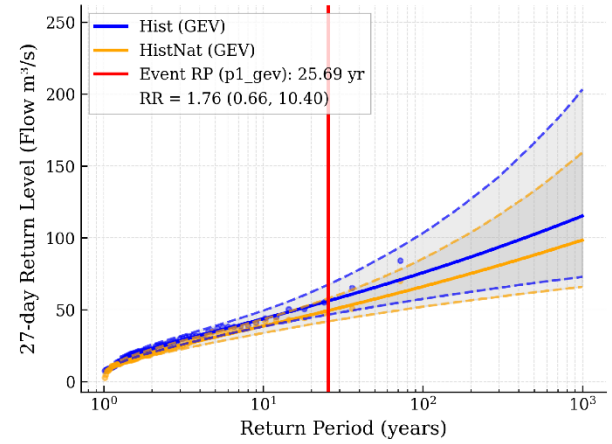
a) BCC-CSM2-MR**b) CNRM-CM6-1****c) IPSL-CM6A-LR****d) NorESM2-LM**

Figure C6: As Figure C5 but for the Raven process-based hydrological model forced with each of the four CMIP6 models.

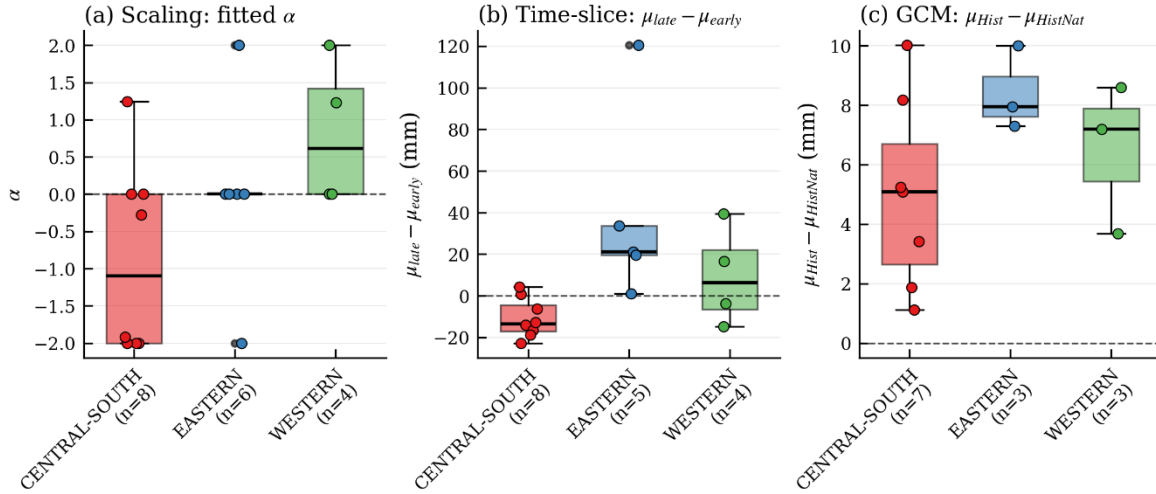


Figure C7: Diagnostic basis for the regional attribution signal in 29-day annual maximum precipitation. Panels show the distribution of three method-specific shifts in the GEV location parameter, aggregated by sub-region: (a) fitted scaling coefficient α from the nonstationary GEV; (b) empirical shift between the late-period and early-period location parameters from the time-slice partitioning, $\mu_{late} - \mu_{early}$ (mm); and (c) CMIP6 multi-model forced response, $\mu_{Hist} - \mu_{HistNat}$ (mm). In each panel, individual stations are plotted as filled circles overlaid on a box plot of the sub-regional distribution. The dashed horizontal line indicates the no-change reference value, and positive values indicate that 29-day precipitation extremes are larger in the warmer climate state.

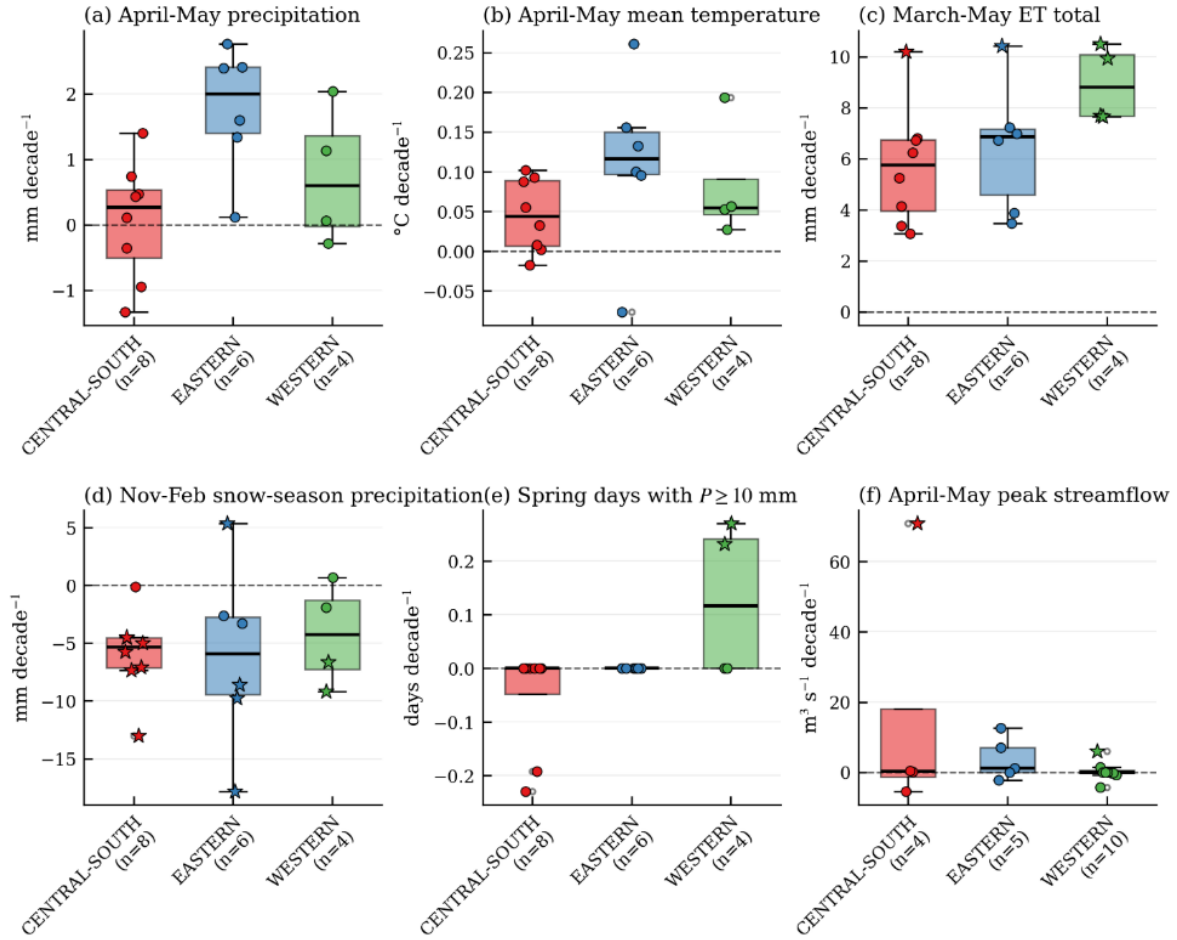


Figure C8: Sub-regional Theil-Sen trends in physical-driver variables for the 1950 to 2022 period. Each panel depicts the distribution of station-level slopes (per decade) for one variable, grouped by sub-region (central-south, eastern, western). Filled circles indicate individual stations and stars highlight stations with a Mann-Kendall p value below 0.05. Panels are: (a) April-May precipitation total, (b) April-May mean temperature, (c) March-May evapotranspiration total, (d) November-February snow-season precipitation, (e) number of spring days with daily precipitation ≥ 10 mm, and (f) April-May peak streamflow at Water Survey of Canada gauges.

The dashed horizontal line in each panel indicates the no-trend reference value.

Appendix D: Additional Copula Diagnostics

Table D1: Best-fitting marginal distributions and copula families by station and period, from the exploratory multi-station screening. Of the stations listed, only Indian Bay is used in the compound snow-temperature attribution reported in the main text (Chapter 4, Section 4.4.6); Brandon A, Great Falls, and Kenora A are retained here only to document the distribution and copula selection across the wider screening set and are not part of the main attribution.

Station	Period	Temperature marginal	Snow marginal	Copula
Brandon A	Early	Student t	Weibull	Gaussian
	Late			
Great Falls	Early	Student t	Weibull	Student
	Late	Normal		
Indian Bay	Early	Normal	Weibull	Frank
	Late	Student t	Gamma	Gaussian
Kenora A	Early	Normal	Gamma	Frank
	Late		Lognormal	Gaussian