The influence of snow initial conditions on ensemble flood forecasting for the Bow river

Bachelor thesis

Isabelle Schippers | s2003805

Date: 21/10/2021

Supervisors:

Dr. ir. Martijn Booij | University of Twente

Dr. Louise Arnal | University of Saskatchewan

Dr. Shervan Gharari | University of Saskatchewan

Dr. ir. Wouter Knoben | University of Saskatchewan

UNIVERSITY OF TWENTE.

Preface

The research carried out in this report is part of my bachelor study Civil Engineering at the University of Twente. During this bachelor thesis I have learned a lot about (ensemble) flood forecasting, assessing those forecasts and modelling with SUMMA, mizuRoute and python. Besides that, I have gained insights in how snow is included in the hydrological cycle and how different floods are treated in Canada compared to the Netherlands.

The project is commissioned by the University of Saskatchewan and in specific its Centre for Hydrology. This part of the university was founded because increased scientific substantiation of water management was considered needed nationally. Research topics within the Centre for Hydrology include hydrology and environment, water resources and global water futures (University of Saskatchewan, 2021). A substantial part of the work that its computational hydrology group performs is aimed at applications in streamflow forecasting, advancing the representation of hydrologic processes in Earth System models and water security assessments (University of Saskatchewan, 2021).

Unfortunately, due to the measures taken against the coronavirus, the entire research has been carried out from home. Nevertheless, the commissioning party was very welcoming and supportive during the research. A special thank you goes to Louise Arnal, Shervan Gharari and Wouter Knoben for guiding me throughout my research. All three have contributed to providing me with the model set-up, given me feedback and a critical look on both the process and the results. Besides that, Louise helped a great deal in understanding the concepts of (assessing the quality of the) forecasting, Shervan has provided the historic measured data and WRF data and Wouter assisted with (understanding the concepts of) modelling with SUMMA and python. I would also like to thank Martijn Booij for his guidance and feedback both during the preparation as the execution of this bachelor thesis. Without all those people this research would not have been possible.

I hope you enjoy reading this report. If you have any questions and/or remarks about the report, they can be send to i.schippers@student.utwente.nl.

Summary

Many parts of the world experience flooding, which can have catastrophic results. Flood forecasts give insights in the probability that such a flooding will occur and because of that one is able to take measures to reduce streamflow or diminish the impacts. However, there is still a lot of uncertainty in those flood forecasts. In cold-region mountain basins it is unknown how much snow-melt contributes to flooding. This is partially a result of the uncertain conditions of the snowpack, which in models translates to the quality of the snow initial conditions and the ability to model the physical processes in those areas accurately.

During the period of 19 to 21 June 2013 intense rainfall and rapid snowmelt resulted in flooding in the Canadian Rocky Mountains and its downstream areas. The flood caused five casualties, monetary damage of approximately six billion Canadian dollars and 200.000 people to evacuate their homes. It is hard to estimate the contribution of the rain-on-snow mechanism, which could potentially have large impacts on floods caused by heavy rainfall and rapid snowmelt. Earlier research showed that especially improved predictions upstream of Calgary could decrease the damages resulting from flooding, like the 2013 Alberta flood.

The research objective for this research is to assess the influence of snow initial conditions on ensemble flood forecasts for different lead times for the Bow river by simulating the 2013 Alberta flood. This is divided into two steps, namely; investigating what the hydrological differences between the different sub-catchments in the study area are and determining how adjusting the lead time affects the influence of snow initial conditions on ensemble flood forecasting for the Bow river during each day of the 2013 Alberta flood.

Both parts of this research employ Structure for Unifying Multiple Modelling Alternatives (SUMMA) for modelling the hydrological processes in the study area and the hill-slope routing and mizuRoute for the routing of the river network.

The hydrological differences between the sub-catchments are assessed based on the snow water equivalent, the precipitation and the streamflow. The snow water equivalent and precipitation, in the hydrological year in which June 2013 falls, are compared with climatology. Furthermore, representative sub-catchments are selected, based on soil type, elevation and land cover, to assess if there is a certain type of sub-catchments that acts different during the 2013 flood, compared to earlier years. Besides that, the value for each of the variables a few days before, during and after the flood are plotted into maps to visually assess the differences between the sub-catchments.

From snow water equivalent analysis it shows that the days before the flooding rapid snow melt occurs in the most upstream areas of the upper Bow. When comparing the hydrological year from September 2012 till September 2013 with previous hydrological years, starting from September 2001, it seems that the snow water equivalent is not necessary higher than previous years, but that the snow melt starts earlier in the year and goes more rapidly. In the same days as the rapid snow melt, heavy precipitation occurs in the front ranges of the study area. This suggests that the important factor in the unfolding of the 2013 flood event was the timing of snowmelt and precipitation.

For each day of the simulated flood curve (22nd of June – 26th of June) flood hindcasts are issued, with lead times ranging from 1 day till 8 weeks. Those forecasts are assessed both qualitatively and quantitively. A qualitative assessment of the overall capabilities of the flood forecast is performed by visually comparing all the different forecasts that are made, with each other and with historic

discharges. The quantitative assessment is performed using the Continuous Ranked Probability Skill Score, Dichotomous skill scores and the reliability diagram. For each method only the streamflow is assessed.

When looking at the flood forecasts for different lead times and different days of the flood peak, it stands out that the flood forecasts only scores well when the forecast is initialised one day before the simulated flood starts. This could mean that the flood forecasts only becomes better because of the improved accuracy of the flow in the river channel. However, a slight differentiation in quality of the flood forecasts can be seen in the river segments that contain the streamflow that resulted from the rapid snowmelt and river segments that contain the streamflow that resulted from heavy precipitation. Therefore, the expectation is that the conditions of the snowpack do have influence on ensemble flood forecasting, however, little. Further research needs to be performed to assess if this improved quality of the flood forecasts for sub-catchments that contain streamflow resulting from rapid snowmelt is indeed caused by the improved snow initial conditions or by different factors.

Table of contents

Table of figures

Table of Tables

1. Introduction

1.1. Research motivation

Many parts of the world experience flooding, which can have catastrophic results. Flood forecasts give insights in the probability that such a flooding will occur and because of that one is able to take measures to reduce streamflow or diminish the impacts. For example, controlled spilling of water or evacuation of certain areas. However, there are still a lot of uncertainties accompanied with flood forecasting.

In cold-region mountain basins it remains unknown how much snow-melt contributes to flooding and in particular when flood forecasts are issued. This is the result of the large uncertainty in snow initial conditions (Vionnet, et al., 2020). Because of the lack of alpine snow measurements, the conditions of the snowpack are uncertain, which not only poses problems for obtaining accurate initial conditions but also makes it more difficult to estimate what physical processes, like the rain-on-snow mechanism, will occur and what their potential contribution to flooding is (Pomeroy, et al., 2016). Research showed that, prediction of snowmelt rate, timing and duration improves when a better snow cover distribution is acquired (Dornes, et al., 2008). Because of the lack of good representation of the snow melt, flood predictions might be largely underestimated, resulting in large damages.

During the period of 19 to 21 June 2013 intense rainfall and rapid snowmelt resulted in flooding in the Canadian Rocky Mountains and its downstream areas. The storm covered a large part of the Bow, Oldman and Elk river basins, visualised in Figure 1, and after the first day the water storage capacity of the rocky soils was filled (Pomeroy, et al., 2016).

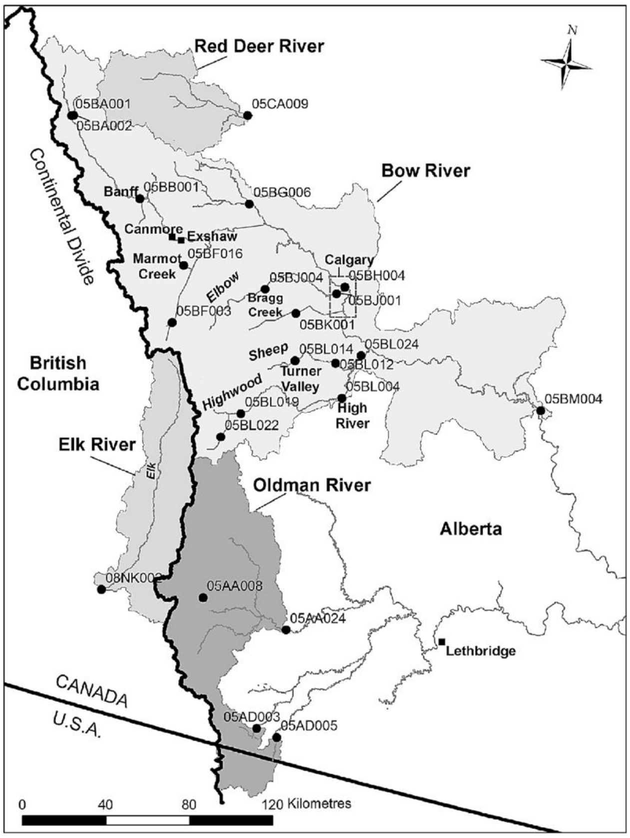

Figure 1 - River basins in Alberta and British Columbia, Canada (Pomeroy, et al., 2016)

Within half a day the discharge of the Bow river increased from 200 m^3s^{-1} to approximately 1700 m^3s^{-1} (Milrad, et al., 2015). Normal seasonal river flows lay between 70 m^3s^{-1} and 400 m^3s^{-1} and the chance of the 2013 Alberta flood is once every 98 years (The city of Calgary, 2021). The return period calculation for this event is however prone to great uncertainty (Pomeroy, et al., 2016). The flood caused five casualties, monetary damage of approximately six billion Canadian dollars and 200.000 people to evacuate their homes (Vionnet, et al., 2020). The damage that large areas suffered was mostly caused by the rapid increase of downhill-moving streamflow, and not so much of local precipitation. Heavy rainfall was forecasted to occur in Southern Alberta during this period, however, the forecasts largely underestimated the most extreme rainfall of this event (Milrad, et al., 2015). It is hard to estimate the contribution of the rain-on-snow mechanism, which could potentially have large impacts on floods caused by heavy rainfall and rapid snowmelt, at large scale, resulting from a lack of alpine snow measurements (Pomeroy, et al., 2016). The largest inaccuracy in the flood forecasting appeared in the area upstream from Banff and the higher elevations of the front ranges. Improved prediction, including improved weather forecasts and accurate streamflow forecasts, upstream of Calgary could decrease exposure to and damage from floods, since it permits short term adaption like evacuation and managing reservoirs (Pomeroy, et al. 2016).

One of the uncertainties in hydrological simulations and forecasts of flood events in complex terrain are the initial conditions. In seasonally snow-covered basins, the uncertain conditions of the snowpack before flooding could result in uncertainties in flood forecasts (Vionnet, et al., 2020). One source of this uncertainty is the displacement of snow through wind, after measurements have taken place (Pomeroy, et al., 2016). Another source of uncertainty, soil moisture conditions, are in spring and summer also dependent on snowpack conditions, because of the snowmelt. Peak flow and flood volume forecasts were highly underestimated when the simulations started with almost no snowpack as initial conditions. In other cases, where initial conditions have a substantial snowpack and coverage in high elevations flood discharge volumes were consistently overestimated. This shows the urge to obtain more accurate snow information in complex terrain (Vionnet, et al., 2020).

1.2. Flood forecasting

1.2.1. Different lead times and frequencies

Flood forecasting can be performed for different lead times and different frequencies. The lead time is the time that passes between the issue date of forecast and the moment for which the streamflow is forecasted. During this lead time the following stages often occur; notification, decision making, warning and action. In case of a hindcast, the lead time can be considered as the time that passes between the date at which the hindcasts is started till the date at which the flood event happens. A hindcast can be described as; one predicts for a date in the past. When generating a flood forecasts/hindcasts it can be initialised at different frequencies (e.g. every day or every month), which one can call the frequency of the forecast.

1.2.2. Ensemble forecasting

In ensemble forecasts one runs the model multiple times with slightly different conditions, instead of only running the most likely outcome like in deterministic forecasting (World Meteorological Organization, 2012). Many hydrological forecast systems make use of lumped and deterministic hydrological models, however, distributed hydrological models and ensemble forecasting have gained serious momentum (Rakovec, 2014). Causes of this shift include, the increase of numerical meteorological data, extension of large capacity computing and a shift in interest from deterministic to risk-based approaches. Especially for forecasts with lead times longer than two or three days, meteorological forecast input causes the largest uncertainty, with the exception of special circumstances, such as seasonal forecasts. Compared to deterministic approaches for flood forecasting, ensemble flood forecasting is reliable for much longer lead times, because of the insights it gives in the uncertainty of the flood forecasts (Wu, et al., 2020). Flood forecasts using a probabilistic approach thus not only determines the most likely forecast, but also gives insights in the probability of extreme or rare events. Besides that, probabilistic approaches give more consistent results on consecutive days than deterministic approaches (Cloke & Pappenberger, 2009).

1.2.3. Current forecasting strategy

Since the study area is located in Canada, only the current forecasting strategy within Canada will be addressed. Within Canada each province has the freedom to choose its own forecasting strategy (Zahmatkesh, et al., 2019) and thus the strategy might differ among provinces. Since the study area lies mostly in Alberta and partially in British Columbia, these are the provinces that will be focused on.

In both British Columbia and Alberta the most common flood types include, rain-on-snow, snowmelt and heavy rainfall. In Alberta, other common types of flooding are ice jam and riverine flooding. The rain-on-snow mechanism can occur in late-spring in interior areas and mostly happens in the autumn and winter in coastal areas. The annual peak flows in areas that receive substantial snow melt often occur in March to June. Flooding caused by heavy rainfall often occurs mid-May to mid-July, because of the high chances of low pressure fields. Because of those difference in what the province is challenged with in terms of flood (forecasting), different provinces have different approaches. These flood characteristics are included in the way that data is collected, and different modelling options and models are chosen (Zahmatkesh, et al., 2019).

Provinces are able to select their own model set-up and therefore there are differences between the spatial resolution that different forecasting organisations use for their models. In 2019, twenty percent of the models used by forecasting centres across Canada were lumped hydrologic models, seventy percent were semi-distributed hydrologic models and ten percent distributed hydrologic models. For those models the lead time ranges from 6 hours till 10 days and the spatial resolution from 2.5 km to 110 km (Zahmatkesh, et al., 2019). The common initialization frequency of operational flood forecasting systems for seasonal forecasting is one month, this however shows low skill for lead times shorter than one month (Lopez, et al., 2021).

1.3. State of the art

There are some studies that conducted research regarding the influence of snow conditions on flood forecasting in mountainous river basins and the effect of different modelling options regarding snow conditions, which will be addressed in this section.

In 2008 research regarding the effects of spatial aggregation of forcing data and initial conditions on modelling snowmelt was executed. This study focused on the effects of the redistribution of snow by wind, between landscape units, and slope and aspect in snowmelt calculations for landscape units on simulation of snowmelt. The study showed that, in most cases, snow ablation was unsuccessfully described by using aggregated initial conditions, whereas when both snow-cover redistribution and slope and aspect effects were incorporated, the prediction of snowmelt rate, timing and duration improved (Dornes, et al., 2008).

In 2009 a study tested the relative contribution of hydrological initial conditions and atmospheric forcing to errors in seasonal hydrological forecasting. This research showed that the uncertainties caused by initial conditions are higher than the uncertainties caused by atmospheric forcing for short lead times, up to approximately one month. The initial conditions have especially a strong impact on forecasts with a short lead time for larger basins. When the lead time is longer than one month, meteorological forcing data is the bigger source of uncertainty (Li, et al., 2009).

A study conducted in 2020 assessed the factors governing the ability to predict late-spring flooding in cold-region mountain basins. This study focused on three potential sources of uncertainty, namely, the snow and moisture initial conditions, the resolution of atmospheric forcing and the representation of the soil texture. Results showed that the main sources of uncertainty were the snow initial conditions, for half of the headwater basins. This shows that, to be able to provide accurate streamflow forecasts during late-spring floods in cold-region mountain river basins, a better representation of the snow pack should be acquired. In this study a lead time of maximum one day was used (Vionnet, et al., 2020).

Previous research already showed that snow (initial) conditions are a large potential source of uncertainty in forecasting seasonal floods in river basins with mountainous terrain (Vionnet, et al., 2020). However, this research has yet to be done in an ensemble context. Besides that, it remains unclear if the lead time and frequency or the spatial resolution of snow initial conditions has more influence on uncertainties of the flood forecasts.

1.4. Problem statement

Streamflow forecasts in mountainous river basins are uncertain, because of the uncertainties in initial conditions and forcing data and specifically the uncertainties in the representation of the snow pack. Currently, it is not clear what the influence of snow initial conditions on flood forecasting are in an ensemble context. Besides that, it still remains to be determined what the influence of varying the lead time and frequency and spatial resolution of these initial conditions is on the accuracy of the flood forecast.

1.5. Study area

As mentioned in section 1.1., earlier research showed that especially improved predictions upstream of Calgary could decrease the damages resulting from flooding, like the 2013 Alberta flood (Pomeroy, et al., 2016). Therefore, the study area covers the upstream part from Calgary of the Bow river basin. Since the city of Calgary was heavily affected by the flood, it is also included in the study area. An overview of the study area and its characteristics is given in Figure 2. The study area starts upstream at Bow lake, and ends at Carseland, 50 km downstream of Calgary. The main river network of the study area is the Bow river and is joined by the Elbow river and the Sheep river along its trajectory. In Figure 2, the river network is plotted as a function of its upstream area. The rivers also have small tributaries in which the water runs-off to the main river network. The average river width of the Bow river is about 70 m downstream of Banff and 30 m upstream of Banff. The average width of the tributaries is less than a meter (Allen, 2021).

The total Bow river basin covers an area of 26200 km^2, of which about 15600 km^2 is the study area. The total length of the Bow river is 578 km, of which about half lies within the study area (National Resources Canada, 2021). The elevation of the area varies from approximately 3500 meters above mean sea level in the mountain area between Banff and Bow lake to approximately 900 meters above mean sea level in the prairies around Calgary (Allen, 2021), visually represented in the bottom right of Figure 2.

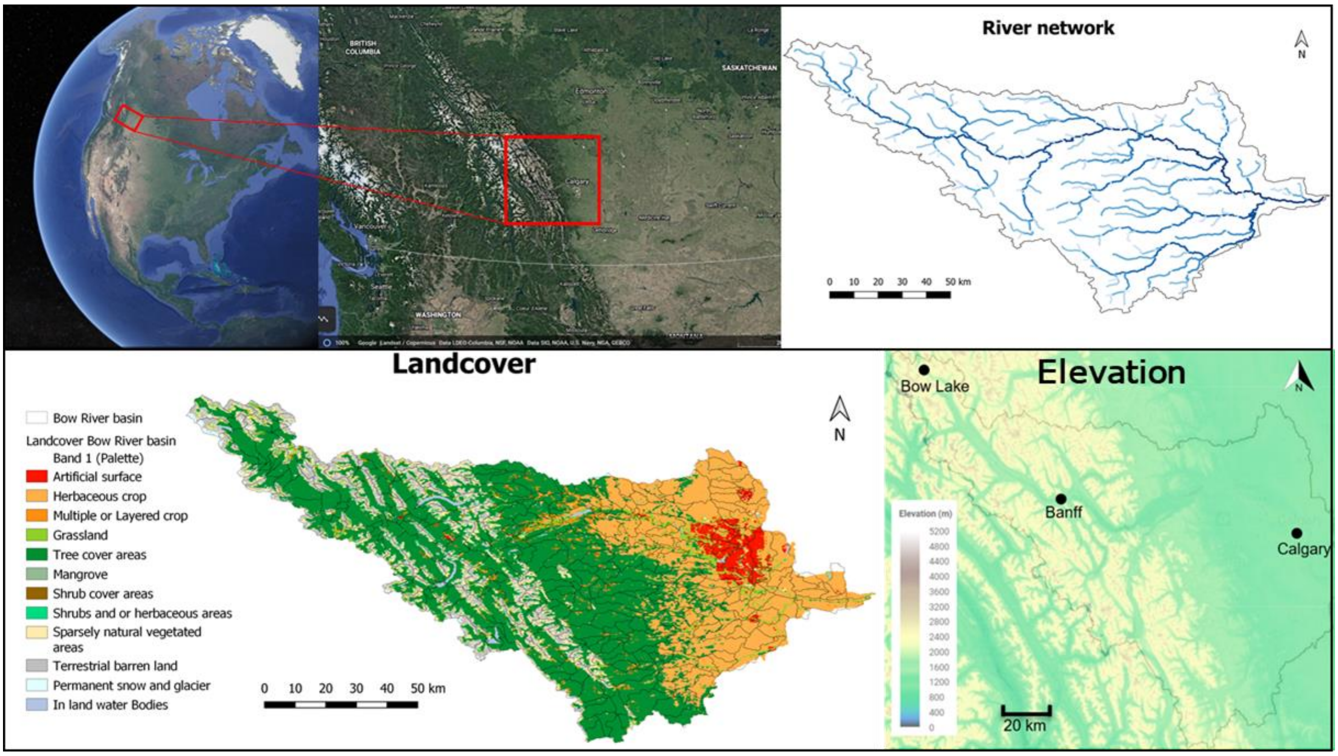

Figure 2 – Topography, landcover, elevation and river network of the study area (Google, 2021) (Allen, 2021) (Bash & Marshall, 2014) (Yamazaki, et al., 2019) (ESA, 2021)

1.6. Research framework

1.6.1. Research objective

The research objective describes the goal that, when achieved, contributes to solving the problem as described in section 1.4. The research objective for this research can be formulated as: assessing the influence of snow initial conditions on ensemble flood forecasts for different lead times for the Bow river by simulating the 2013 Alberta flood.

1.6.2. Research questions

The research can be divided into two large parts, where the first part focusses on understanding the hydrological and thermodynamical conditions that caused the 2013 Alberta flood and the second part focuses on hindcasting the event and determining how the lead time and frequency representation of snow initial conditions affect the flood forecast.

The research objective as described in section 1.6.1. can be divided into the following research questions:

1. *What are the hydrological differences between the different sub-catchments in the study area?*
   To get a better understanding of how the event unfolded and what relevant sub-catchments are to focus on in the second part of this research, it is important to investigate the hydrological differences between the different sub-catchments.

2. *How does adjusting the lead times affect the influence of snow initial conditions on ensemble flood hindcasting for the Bow river during each day of the 2013 Alberta flood?*

As mentioned in section 1.2.3, the current frequency of seasonal flood forecasting is once a month. The accuracy of the forecast might change when the frequency of the forecast is adjusted. How much the accuracy of the forecast changes gives useful insights in if adjusting the frequency of snow initialization is worth the extra amount of computational power needed to perform these forecast. Since this study only focusses on a single event this frequency is largely dependent on changing the lead time of the forecast. However, by looking at each of the days of the flood peak individually, the influence of different starting dates, for equal lead times can still be assessed. Improved streamflow forecasts could decrease exposure to and damage from floods, since it permits short term adaption like evacuation and managing reservoirs (Pomeroy, et al. 2016). In this study the potential contribution of improved representation of the snowpack is determined for different lead times.

1.7. Reading guide

The research in this report is divided into two parts, simulation of the hydrological and thermodynamical processes over the period of October 2000 till the spring of 2013 and the flood hindcasts for the 2013 Alberta flooding. For both parts the same model is used, which is described in section 2.1, along with other foreknowledge used in this research. In section 2.2 the methods that are used to conduct this research are described, organised by research question. In chapter 3 the results can be found. In chapter 4 one is able to find discussion of the used materials, methods, interpretation of the results and limitations of this research. The conclusions and recommendations can be found in chapter 5. This report also contains a number of appendices, in which supporting figures and tables can be found, to which are referenced in the text.

2. Materials & Methods

2.1. Materials

2.1.1. Hydrological model

For all research questions within this research, as described in section 1.6.2, the Structure for Unifying Multiple Modelling Alternatives (SUMMA) and mizuRoute will be used. SUMMA is used to simulate the hydrological and thermodynamical processes that happen in the study area over time. The main characteristic of SUMMA that makes it useful for this research, compared to other runoff models, is that it simulates the physical processes in the sub-catchment detailed and that the model works independently of its location. This is important to accurately model snow and ice processes, especially with the complex terrain of the study area. Since SUMMA does not simulate the routing between different sub-catchments, mizuRoute is used for the routing between the different sub-catchments. Another reason for choice of SUMMA and mizuRoute, instead of other models that could be used for this study, is that those are used by the research group at which this research is executed and thus there can be benefitted from the local expertise with the chosen models.

2.1.1.1. Structure for Unifying Multiple Modelling Alternatives

In this section the Structure for Unifying Multiple Modelling Alternatives, further called SUMMA, is described. SUMMA is based on two propositions. The first is that the majority of hydrologic modellers have a similar understanding of the effect of dominant fluxes of energy and water on the time evolution of hydrologic and thermodynamic states, but that there is uncertainty about the most correct way to implement fluxes as equations. SUMMA makes it easy to switch between different equations for a given process. The second proposition is that spatial variability and hydrologic connectivity within the model domain are the major scientific issues in hydrological model development. The model domain of SUMMA includes the area between the river channel and the atmosphere above vegetation canopy. Within its model domain, SUMMA simulates both hydrological and thermodynamic processes (Clark, et al., 2015a). The different processes conceptualised in the model are visualised in Figure 3.

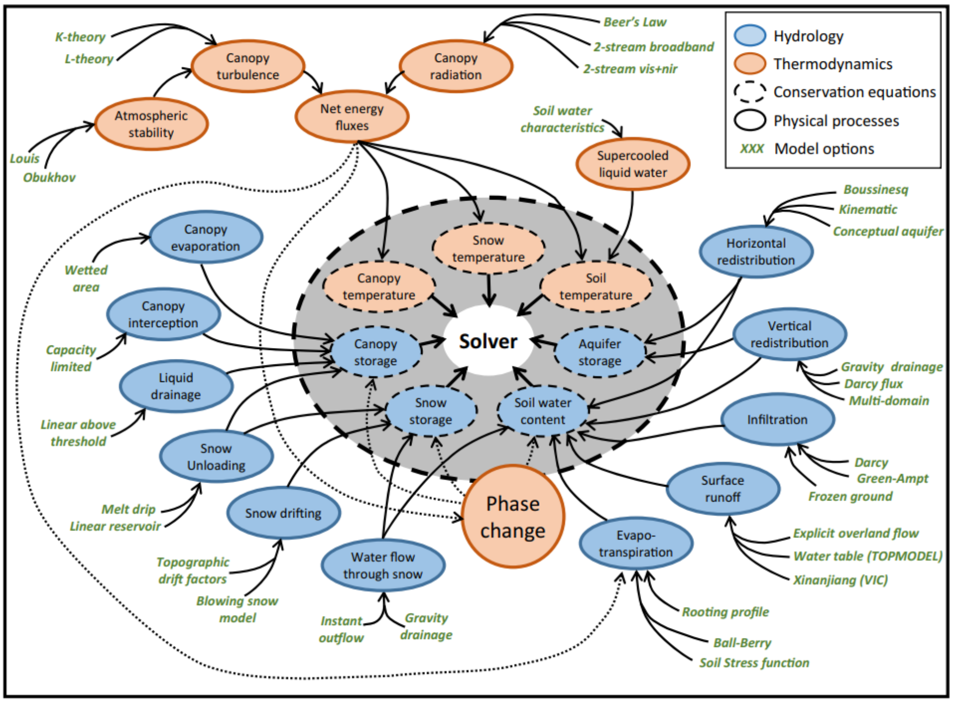

Figure 3 - Conceptual model SUMMA (Clark, et al., 2015a)

The way SUMMA is organised offers possibilities to employ a flexible hierarchical spatial structure. This hierarchy is made up from grouped response units (GRUs) and within each GRU hydrological response units (HRUs). Some of the key characteristics of the HRUs are that they do not require to be spatially contiguous and can be of any size and shape (Clark, et al., 2015a). Compared to other modelling frameworks, SUMMA progresses in systematic analysis of competing modelling choices, regarding both spatial discretization and process formulation and parametrization. This supports research on how the choice of the spatial discretization approach affects basin-wide runoff and evapotranspiration fluxes, by supporting multiple modelling options for spatial variability and hydrological connectivity (Clark, et al. 2015b). Both lumped hydrologic models and a wide range of spatially distributed models can be implemented with SUMMA (Clark, et al., 2015a).

Input that SUMMA uses are, NOAH-MP tables, which overwrite the default soil and vegetation parameters, the topology of the study area, meteorological forcing data, local attributes, local parameters and basin parameters. Local attributes are hydrological response unit (hru) specific and include, among other things, the elevation, longitude, latitude and surface area of the hru. The local parameters specify spatially constant parameter values for different parameter within SUMMA, including a upper and lower bound for that parameter (Clark, et al., 2015c). Input that describes how the model should perform and other technical details are explained in Appendix A.

2.1.1.2. mizuRoute

The water flow between the different sub-catchments is routed by using mizuRoute. The mizuRoute tool processes the runoff of each element of a spatially distributed model, creating spatially distributed streamflow. The mizuRoute tool works by first using gamma distributions to estimate the temporal delay in runoff within a certain sub-catchment (hill-slope routing) and after that the river network is routed. The hill-slope routing can also be done by other models and in this study is performed with use of SUMMA. By using mizuRoute, streamflow at any defined spatial point in the model can be obtained (Mizukami, et al., 2016). The technical details of how mizuRoute is used in this research can be found in Appendix A.

2.1.1.3. Model set-up

The model set-up, as described in this section, was provided by Louise Arnal, Shervan Gharari and Wouter Knoben from the Canmore Coldwater Lab, part of the Centre of Hydrology, at the University of Saskatchewan. The model set-up contains all needed data to run simulations from October 2001 till October 2013 and has been calibrated already.

As discussed in section 2.1.1.1, SUMMA supports multiple modelling choices, and thus the modelling choices need to be defined. In Table 3 in Appendix B there is given an overview of the modelling choices that are used within the model set-up. The modelling decisions represent a set of standard decisions used for the study area in this research within the Centre of Hydrology of the University of Saskatchewan and are based on their experiences with SUMMA for this modelling domain.

The model is forced with data from Weather Research and Forecasting, further called WRF, meteorological reanalysis. The WRF model is a numerical weather prediction model partially designed for (operational) forecasting systems. It combines conventional precipitation and surface and upper-air radar data with satellite data (Powers, et al., 2017). The specific dataset that is used in this research is 2000-2013 WRF simulation that covers large part of North America at 4 km grid spacing (Rasmussen & Liu, 2021). In Table 4 in Appendix B there is given an overview of the used variables in this dataset. The advantage of using this dataset, compared to observed data, is that the data is available over the whole domain, instead of a few fixed points and still has a high resolution (Powers, et al., 2017).

Topographic properties of the study area have to be coupled to the model. This is done by making use of the Multi-Error-Removed-Improved-Terrain, further called MERIT, Hydro catchment delineation. MERIT Hydro is a high resolution global flow direction map that combines water body data sets with elevation data (Yamazaki, et al., 2019). From this, the river flowlines and sub-catchments are vectorised (Lin, et al., 2019). In Figure 4 an overview of the modelling domain is given.

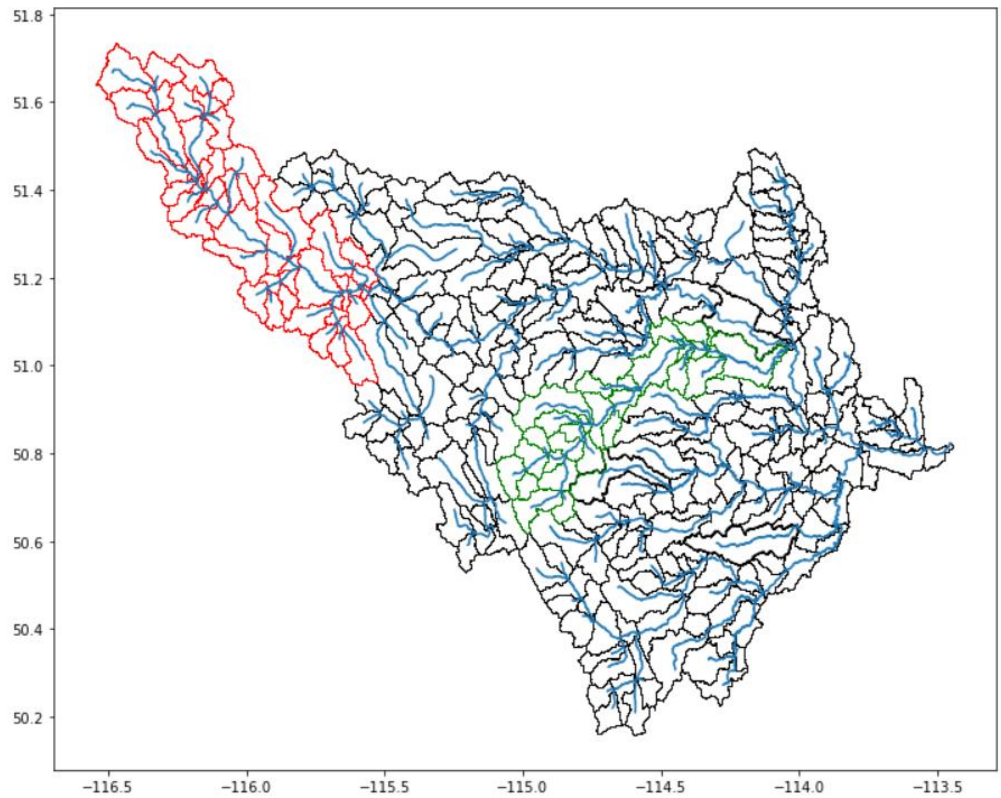

Figure 4 - Modelling domain with upper Bow (red) and Elbow (green) sub-catchments

The model has been calibrated at three different locations in the modelling domain, the Bow at Banff (upper Bow), the Elbow at Sarcee bridge and the Bow river at the Carseland dam (outlet of study area). This results in three different parameter sets, one for the sub-catchments within upper Bow, one for the sub-catchments within Elbow and one for the other sub-catchments in the study area. Because of that, the diverse landscape and thus different characteristics within the whole study area are better represented.

For each of the three locations the optimal parameters were determined with an dynamic dimensioned search tool (OSTRICH, 2017), which is a calibration algorithm designed for models with many parameters and is ideally suited for models which require high computational power (Tolson & Shoemaker, 2007), like SUMMA. This is performed with the physically possible ranges of the parameters, found in the basin parameter file and local parameter file. The parameters were selected based on the Kling-Gupta efficiency. In Table B-3 in Appendix B, an overview of parameters, for which calibration was performed, is displayed.

The calibration was performed with data from October 2002 until October 2008 and was validated with data from October 2000 until October 2012. Note that, until the spring of 2006, there are no measurements during the winter period.

The assessment metric that was used, the Kling-Gupta efficiency (KGE), assesses the difference between forecasted and observed data. The range of the Kling-Gupta efficiency is -∞ to 1, where 1 describes a perfect fit (Gupta, et al., 2009). When the value of the Kling-Gupta efficiency drops below -0,41, it means the average value gives a better prediction than the model (Knoben, et al., 2019).

After calibration the following scores for the KGE were obtained; a KGE of 0.82 for the Bow river at Banff, a KGE of 0.72 for the Elbow river at Sarcee Bridge and a KGE of 0.76 for the Bow river at the Carseland dam.

2.1.2. Model and input evaluation

2.1.2.1. Model evaluation

The model set-up is calibrated for streamflow only and only until 2008. In this section the model is evaluated by comparing the streamflow measurements with streamflow simulations, for the year 2013 and by comparing snow water equivalent (SWE) measurements with snow water equivalent model output.

2.1.2.1.1. Streamflow evaluation

In Figure C-7, Figure C-8 and Figure C-9 in Appendix C, the measured and simulated streamflow are plotted for the period of 2001-2013 for the Elbow, the Bow at Banff and the Carseland dam, respectively. When comparing 2013 with previous years, it stands out that the simulation for the outlet is quite well and for the Bow at Banff also is quite reasonable, in line with previous years. However, the streamflow simulation for the Elbow is wildly underestimated, where in other years the peaks of the simulation and the measured discharge match quite well. This could be because the streamflow for the Elbow is exceptionally high in 2013. Therefore one needs to be careful when drawing conclusions. When taking a more in depth look at the 2013 flood, a delay for the simulated flood curve can be observed, as visualised in Figure C-10, Figure C-11 and Figure C-12 in Appendix C.

2.1.2.1.2. Snow water equivalent evaluation

Because of the limited amount of snow observation stations in the study area, they are all selected for the validation. In Figure 5, an overview of the locations of the snow observation stations is given and in Table C-1 in Appendix C the characteristics of those snow stations are displayed. Snow observation data has been retrieved from ECCC by Louise Arnal and is a revised version from the data set; Canadian historical snow survey data (Government of Canada, 2021).

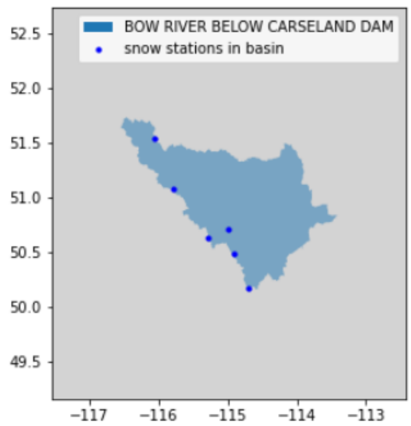

Figure 5 – Six snow observation stations

When comparing the measured SWE with the simulated SWE for each of the snow observation stations and their corresponding hydrological response unit, it stands out that the difference is small for cases that the mean elevation of the hydrological response unit is substantially higher than the elevation at the snow observation stations. The difference is larger in cases that the mean elevation of the hydrological response unit in which the snow observations station is located is similar to or smaller than the elevation of the snow observation station. This suggests that the SWE is systematically underestimated. This is visually displayed in Figures C-1 up and until C-6 in Appendix C.

2.1.2.2. Input evaluation

Since literature shows that the precipitation is an important contributor to the 2013 Alberta floods (Pomeroy, et al., 2016) (Milrad, et al., 2015), the precipitation input data is validated. This validation is performed by comparing measured precipitation at five different locations in the study area with average of the simulated precipitation of the corresponding sub-catchment. Locations are selected based on data availability during the simulation period from 2000 up and until 2013 and proximity to other selected observation stations.

The simulated and measured precipitation are plotted for each of the five locations and can be found in Figure C-7 up and until C-11 in Appendix C. When comparing the measured precipitation with the precipitation data that is used as input in the model, as described in section 2.1.1.3. it stands out that the two are more similar when the elevation of that location is lower. This seems logical since at lower elevations the precipitation varies less throughout the sub-catchment (Shaw, et al., 2011) and the input data is based on multiple sources that are not dependent on one fixed location and the validation data is location dependent.

2.1.3. Extended Streamflow Prediction (ESP) forecast

The forecasting method that will be used in this research is the Extended Streamflow Prediction, further called ESP. ESP is designed for water supply forecasting in regions with snowmelt and can also be used to predict spring floods (Day, 1985). These factors make the forecasting method fit for this study. ESP employs a hydrologic model to predict future streamflow. Current conditions of soil, snow, moisture and river are forced with historic meteorological data. The separate years of the meteorological data are considered as possible representation of the future and will be a separate ensemble member in the forecast (Day, 1985). In Figure 6 a visual representation is given of how this type of forecast looks, in which initial conditions is shortened to ICs.

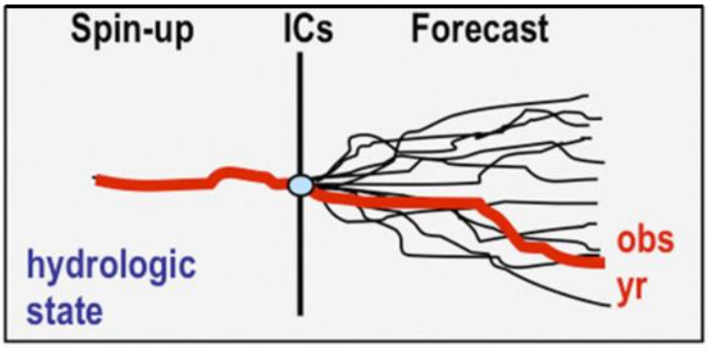

Figure 6 – ESP (Wood, et al., 2016) Each ensemble forecast member is generated from forcing data observed during the forecast period but in different years.

2.1.4. Criteria for assessing the flood forecasting quality

Flood forecasts have different aspects on which they can be assessed and thus different assessment criteria evaluate a different part of the forecasting process. The overall forecasting quality can be assessed by using the continuous ranked probability (skill) score, the resolution of the hindcasts is determined with the ROC-diagram and the reliability with the reliability diagram. By including a histogram of the sample size in the reliability diagram, the sharpness is also assessed. An advantages of using both the ROC-diagram and the reliability diagram is that they complement each other well, since the ROC-diagram is conditioned on the observations and the reliability diagram is conditioned on the forecasts (World Weather Research Program, 2021). These criteria will be used in the assessment of the flood hindcasts performed for the second part of the research, as described in section 2.2.2.3.

2.1.4.1. Continuous Ranked Probability (Skill) Score

The focus of the continuous ranked probability (skill) score (CRP(S)S) lies on the complete range of a specific parameter and can be interpreted as the integral of the Brier score over all possible threshold values for the concerning variable (Hersbach, 2000). The Brier score is a verification metric in which the quadratic difference between the forecast probability and the observed for each occasion is summed and divided by the total number of occasions (Brier, 1950). The CRPS is calculated with the python package *properscoring* (PyPI, 2021), in which the CRPS is a built-in function. The unit of the CRPS is equal to the unit of the variable for which the CRPS is calculated. This does result in a higher value for the CRPS when the value of the parameter is higher. Therefore, the CRPS can be expressed as a value in which it is compared with the baseline, the CRPSS. In this case, assessments in which the assessed variables have a different order of magnitude can be more easily compared. The perfect score for the CRPSS is 1 (Hersbach, 2000).

2.1.4.2. Dichotomous skill scores

With the dichotomous skill scores can be tested how good an event forecast is. A dichotomous forecasts predicts if an event will happen or not (World Weather Research Program, 2021). To verify the models forecasts, the contingency table as depicted in Table 1 is used. There are four different possible combinations, called the joint distribution. Those combinations are:

- Hit: the event was both forecasted and did occur.
- False alarm: the event was forecasted but did not occur.
- Misses: the event was not forecasted but did occur.
- Correct negatives: the event was not forecasted and did not occur.

Table 1 - Contingency Table (World Weather Research Program, 2021)

		Observation	
		Yes	**No**
Forecast	**Yes**	Hits	False alarms
	No	Misses	Correct negatives

There are different equations, depending on the focus, that can be used to assess the scores of the contingency table (World Weather Research Program, 2021). For this study the probability of detection (POD), also called the hit rate, and the probability of false detection (POFD), also called the false alarm rate, are used. The POD can especially be used well for events that occur with a low frequency. The POD ranges from 0 and 1 and its perfect score is 1. The POD can be calculated with use of equation 1. The POD is very sensitive to the climatological frequency of the event. Besides that, it does not take into account false alarms and therefore should be combined with a metric that does take this into account (World Weather Research Program, 2021).

$$POD = \frac{Hits}{Hits+Misses} \quad \text{Eq. 1}$$

With the POFD it is calculated in how many cases the forecasts predicts a flood that actually does not occur. The POFD has a range from 0 to 1 and its perfect score is 0. The POFD can be calculated with use of equation 2 (World Weather Research Program, 2021).

$$POFD = \frac{False\ alarms}{False\ alarms + Correct\ negatives} \qquad \text{Eq. 2}$$

A common way to assess the quality of the forecast is to discriminate between events and non-events and thus measure the resolution of the forecast is the relative operating characteristic (ROC) diagram. In the ROC diagram the POD is plotted against the POFD for different streamflow occurrence probabilities, as visualised in Figure 7.

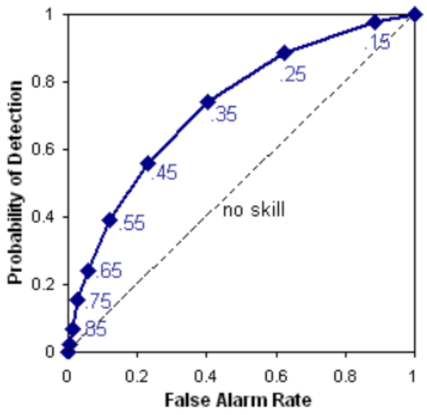

Figure 7 - The Relative Operating Characteristic diagram (World Weather Research Program, 2021)

With the ROC-diagram the ability of the forecasts to discriminate between two different outcomes (the resolution) is assessed. A perfect score follows the line from the bottom left to the top left and then to the top right of Figure 7 (from (0,0), to (0,1) to (1,1)). When the curve is above the diagonal, as in Figure 7, this shows that the forecasts at least have some skill and when the curve follow the diagonal or is below the diagonal, it shows that the forecasts have no skill (World Weather Research Program, 2021).

2.1.4.3. Reliability diagram

The reliability diagram shows the quality of multiple factors of the forecast, the reliability, the resolution and the sharpness (in the histogram). The reliability part assess to which extend the predicted probabilities of an event correspond with the observed frequencies. When the model is calibrated well, the perfect reliability follows the 1:1 diagonal. When the line is parallel to the perfect reliability but higher, there is under-forecasted and when the line is parallel to the perfect reliability but lower, there is over-forecasted. When the line is more horizontal, this shows a poor resolution In case of under-forecasting or over-forecasting the forecast can be improved by calibration (World Weather Research Program, 2021).

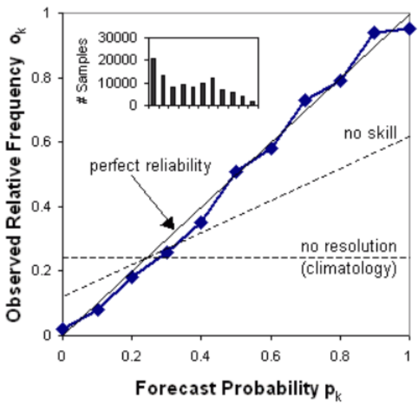

Figure 8 - Reliability diagram (World Weather Research Program, 2021)

The sample size in each probability bin is included in the reliability diagram in the form of a histogram. With this histogram the sharpness of the forecasts can be assessed. Sharpness can be defined as a tilt to forecasting values near 0 or 1, instead of values clustered around the mean (Ranjan, 2009) (World Weather Research Program, 2021). When the forecast is sharp, the histogram should be U-shaped (Ranjan, 2009).

2.2. Methods

In this section the methods used within this research are explained for each sub-question, separately. An overview of the methodology for the entire research is visualised in Figure 9. In the figure, the yellow boxes represent input (for a different sub-question), the red boxes represent results for each of the sub-questions and the blue boxes represent the steps that need to be taken in between input and output.

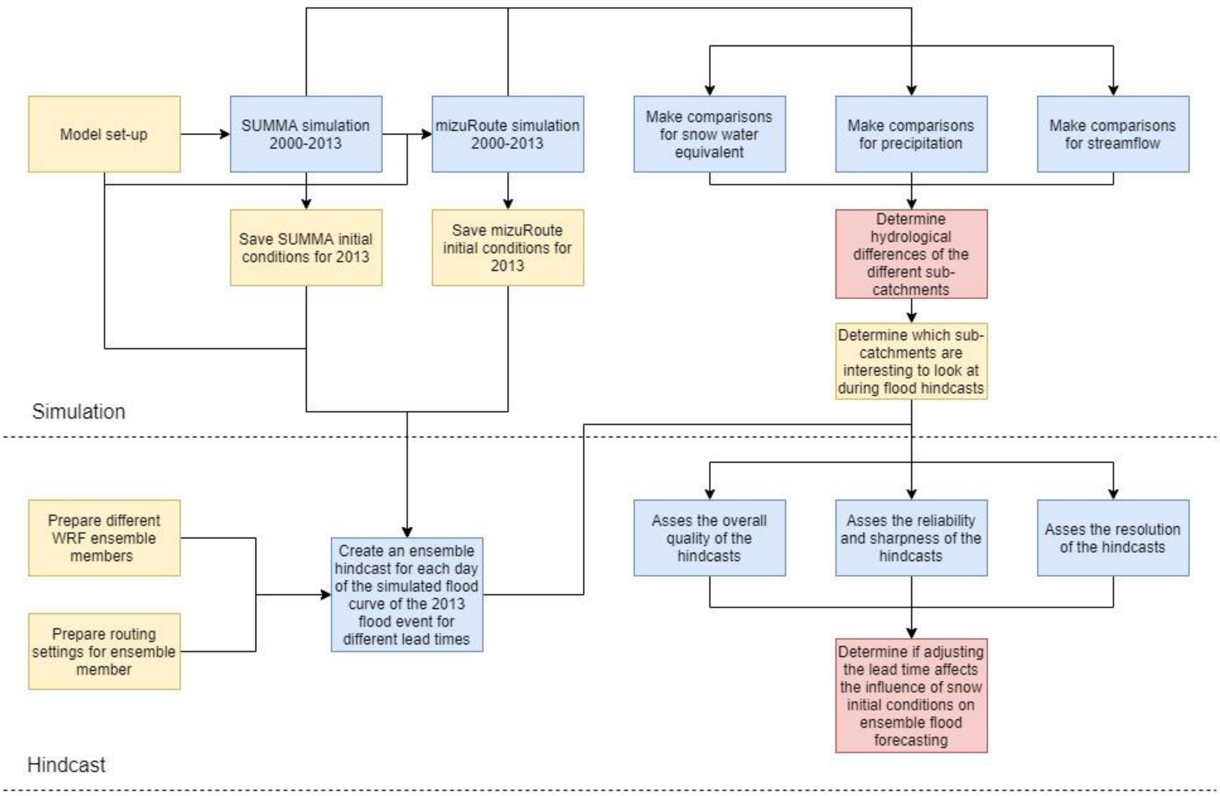

Figure 9 - Overview of methodology

2.2.1. Hydrological differences between (representative) sub-catchments

The hydrological differences between the different sub-catchments in the study area are investigated by simulating the hydrological and thermodynamic processes in the study area from 2000 to 2013 and looking at the different properties of the sub-catchments after the warm-up period. There is opted for a warm-up period of 1 year. There will especially be looked at the contribution of each of the sub-catchments to the total streamflow, the precipitation within each of the sub-catchments and the snow water equivalent of each of the sub-catchments.

Of the three variables that are studied, the method of research for the snow water equivalent [kg m^{-2}] and precipitation [mm] are the same. First, the values of those simulated variables for different representative sub-catchments are compared for the entire period of the simulation. The selection of different sub-catchments is made based on diversity in landscape, soil type and elevation, by selecting a sub-catchment for each possible combination of those aspects. An overview of the characteristics of the selected sub-catchments is depicted in Table 2.

Table 2 - Selected representative sub-catchments for comparison of different hydrological conditions and their characteristics

HRU number	Elevation	Soil type	Land-use	Surface area [km^2]
71038127	High	Sandy Loam	Terrestrial barren land	33.9
71032292	High	Loam	Terrestrial barren land	25.6
71028377	High	Loam	Tree cover areas	108.3
71029721	Medium	Loam	Tree cover areas	50.5
71034018	Low	Loam	Artificial surface	104.9
71028976	Low	Clay Loam	Artificial surface	6.9
71030555	Low	Loam	Herbaceous crop	35.6
71039072	Low	Clay Loam	Herbaceous crop	62.9
71028014	Low	Loam	Grassland	18.2

This step is executed to see which type of sub-catchments potentially contribute a lot at the time of the flooding. Secondly, the hydrological year in which the June 2013 flood lies is compared with the previous hydrological years, starting after the warm-up period, for both the snow water equivalent and the precipitation. Since the built-up of the snowpack starts in September, the used hydrological year runs from the first of September till the 31st of August. For this comparison the median value of the different sub-catchments is used and for the precipitation additionally the mean value, since if there are a lot of dry years the median might give a wrong representation. Furthermore, the values for the snow water equivalent and precipitation a few days before the event, the days of the event and a few days after the event of each sub-catchment are plotted into maps. The resulting maps will show the conditions under which flooding occurred and thus which sub-catchments might potentially have a lot of influence and which potentially have little influence on when flooding occurs.

The streamflow is analysed by plotting the streamflow for a few days before the event, the days of the event and a few days after the event of each river segment in maps. This is done to look which river segments had significantly a higher streamflow than on other days. After that, the percentage of snow melt of the total runoff and the percentage of precipitation of the total runoff is determined for each sub-catchment and plotted in a map. This will give insights in how much the snow water equivalent and the precipitation contribute to the increased water levels.

The outcomes of the streamflow analysis and the results of each of the three steps for the snow water equivalent and precipitation are then compared to determine hydrological differences between the sub-catchments and which sub-catchments had a lot of influence in the unfolding of the 2013 flood event and which had little influence in the unfolding of the 2013 flood event.

2.2.2. How do different lead times affect the influence of snow initial conditions on ensemble flood forecasting for the Bow river during the 2013 Alberta flood?

2.2.2.1. Initial conditions

2.2.2.1.1. SUMMA initial conditions

The warm-up period for the simulation of streamflow with the model set-up as described in section 2.1.1.3, is longer than the lead times used for (seasonal) flood forecasting. Therefore, the conditions of the snowpack need to be accurate at the start of the forecast, which is done by giving the model initial conditions. Those initial conditions are obtained from model spin-up. This research focusses on the influence of the snow initial conditions, but in SUMMA it is only possible to restart all initial conditions at the same time. An overview of those initial conditions within the SUMMA model are presented in Table E-1 in Appendix E.

The initial conditions can be obtained by saving the corresponding conditions that the model has at a certain moment in the simulation for 2000-2013. The conditions of each day from the 27th of April 2013 till the 25th of June 2013 will be saved. This is done by first running the whole simulation for 2000 till the 27th of April 2013 without specifying initial conditions and then saving an initial conditions file at the end of the simulation. After that, a new simulation will be executed starting from the 27th of April 2013 with the saved conditions as initial conditions. Initialising the conditions for a simulation makes sure that there is no warm-up period and thus saves a lot of simulation time and memory storage. The new simulation will be performed until the 25th of June 2013 and conditions will be saved for each day.

2.2.2.1.2. Routing initial conditions

To be able to perform the flood hindcasts there needs to be an accurate representation of the runoff at each specified location in the model domain at the start of the flood forecast. Those conditions of the flow network can be obtained by running a simulation up and until the moment that the flood hindcast starts. Those conditions can be saved by adjusting the mizuRoute control file. The conditions will be saved on the dates that accord with the different lead times used as described in section 2.2.2.1.1.

2.2.2.2. Hindcasting

For each day of the flood curve hindcasts are be performed, which gives insights in which parts of the flood curve are forecasted well and which not so much. Since the simulated flood happens from the 22nd of June until the 26th of June, as stated in section 2.1.2.1.1, those dates are used, instead of the actual dates that the 2013 Alberta flood happened (19th – 21st June). Flood hindcasting will be done for different lead times, ranging from one day to eight weeks. This includes the following lead times; one day, two days, three days, four days, five days, six days, one week, two weeks, three weeks, four weeks and eight weeks. To perform those hindcasts, the initial conditions as explained in section 2.2.2.1 are used. However, because of a bug in SUMMA, one is not able to start a simulation on a day where there is canopy ice and the temperature is also above 0, therefore, some hindcasts are started a day earlier or later. The exact starting days can be found in Table E-2 in Appendix E. Other technical details are explained in Appendix A.

2.2.2.3. Assessing the hindcasts

The flood hindcasts will be assessed based on the streamflow prediction on a qualitative basis by looking at the different forecasts and on a quantitative basis by the different criteria described in section 2.1.4. In this section the application of those criteria within this research will be described.

Since seasonal floods in the study area generally occur between 15 May and 15 July (The city of Calgary, 2021), the qualitative assessment will be compared with different occurrence frequencies from this period. Those occurrence frequencies and their corresponding streamflow values can be found in Table F-3 in Appendix 3. The period from 15 May until 15 July is also used to determine the low threshold for the reliability diagram.

2.2.2.3.1. Continuous ranked probability (skill) score

The overall quality of the hindcasts for each day of the flood curve and each lead time is assessed with the CRPSS. The CRPS is not used for assessment of the flood hindcasts, but only to calculate the CRPSS, since with this value it is hard to compare different locations and different days of the flood curve. The variable for which the CRPSS is calculated is the streamflow in m^3s^{-1}, this means the unit of the CRPS is m^3s^{-1} too. The ensemble member of 2013 is used to compare with, since that is considered 'the perfect simulation'. There is opted to compare with the 'perfect simulation' instead of observations, because of the measuring inaccuracies that come with observations and to cover for model inaccuracies. This step is performed for the locations that are selected in the first part of this research, namely, Lake Louise, the Bow at Calgary, the Elbow at Calgary, Carseland dam and the Sheep river, as described in section 3.1.4.

2.2.2.3.2. Dichotomous skill scores

Regarding the dichotomous skill scores, the quality of the flood forecast is assessed by plotting the ROC-diagram. Since this part assesses how well the flood forecasts perform in an operational system and because of data availability, this assessment is performed for the Bow at Calgary and the Elbow at Calgary only. Distinction between an event and a non-event is made based on if a certain threshold streamflow is met. The thresholds that are used to determine the dichotomous skill scores are based on different flood impacts. Those flood impacts and their corresponding thresholds for both locations are depicted in Table F-1 in Appendix F. The ensemble member of 2013 is again used to compare with, since that is considered 'the perfect simulation'.

2.2.2.3.3. Reliability diagram

Since the sample size of individual lead times is too small to create a reliability diagram, the reliability of the forecasts is determined for each locations only once. This is done using all lead times for the 26th of June and including all hindcasts issued for each day in the period 12 June until 26 June. The ensemble member of 2013 is again used to compare with, since that is considered 'the perfect simulation'.

The reliability diagram is divided into 5 bins, ranging from 0 to 1 with steps of 0.2. Variations in the occurrence frequency, that is used to determine the low threshold, are used for the different locations, to guarantee that more than one bin was filled. Those percentiles of occurrence and there corresponding thresholds are depicted in Table F-2 in Appendix F.

3. Results

3.1. Hydrological differences between (representative) sub-catchments

The hydrological differences between the sub-catchments are determined based on the precipitation, the snow water equivalent (SWE) and the streamflow within each of the sub-catchments. In each of the paragraphs of this section one of the variables is addressed and after that there is looked at which would be interesting to look at.

3.1.1. Snow water equivalent (SWE)

When comparing the hydrological year from September 2012 till September 2013 with previous hydrological years, starting from September 2001, it seems that the snow water equivalent is not necessary higher than previous years, but that the snow melt starts earlier in the year and proceeds at an increased rate. This is visualised in Figure 10.

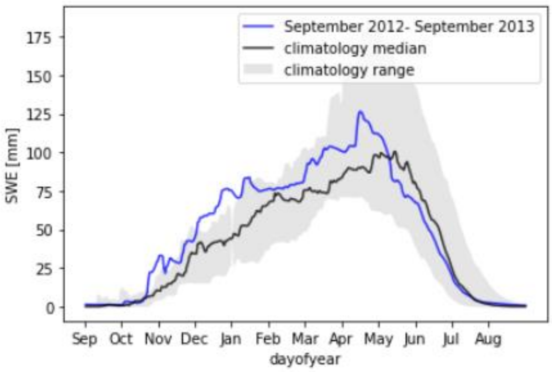

Figure 10 - Comparison of the basin average SWE in 2012-2013 and other hydrological years

That the total amount of snow water equivalent is similar to previous years is also supported by looking at the selected representative sub-catchments, shown in Figure 11. For most of the sub-catchments the snow water equivalent is similar to previous years, only high elevation sub-catchments with the soil type sandy loam have substantially more snow water equivalent than previous years. Note that, in this sub-catchment snow does not completely melt, because there is a glacier in real life. SUMMA does not model glaciers but the weather conditions are as such that the snow does not always melt.

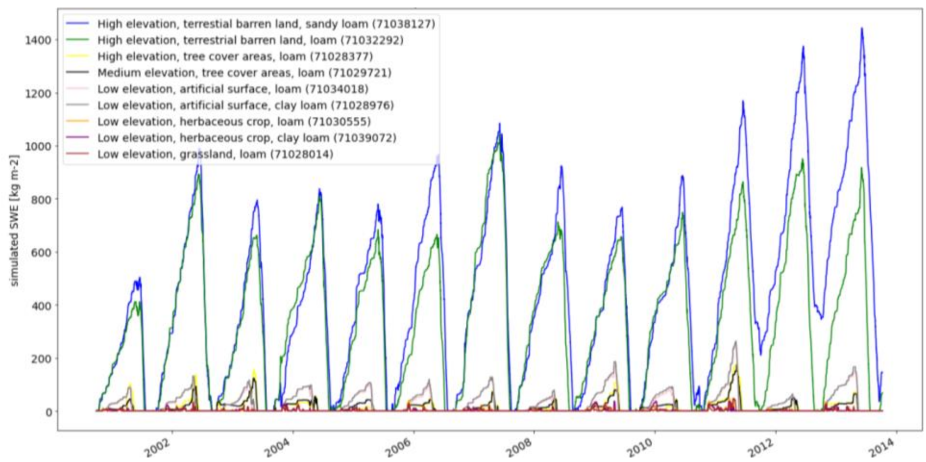

Figure 11 - Simulated snow water equivalent over time for selected sub-catchments

When looking more in depth at each of the sub-catchments during the days before, during and after the 2013 flooding, it shows that, in the days leading up to the flooding there is a sudden increase in the amount of snow melt in the most upstream sub-catchments of the Bow river, as visualised in Figure 12. When comparing this period in 2013 with previous hydrological years, it shows that for those sub-catchments the snow water equivalent was higher than average before the flood and average or below average during and after the flood, which is visualised in Figure G-2 in Appendix G. This suggests that the most upstream catchments of the Bow river had a significant influence on the occurrence of the flooding, because of their contribution to the streamflow resulting from the meltwater of the snow.

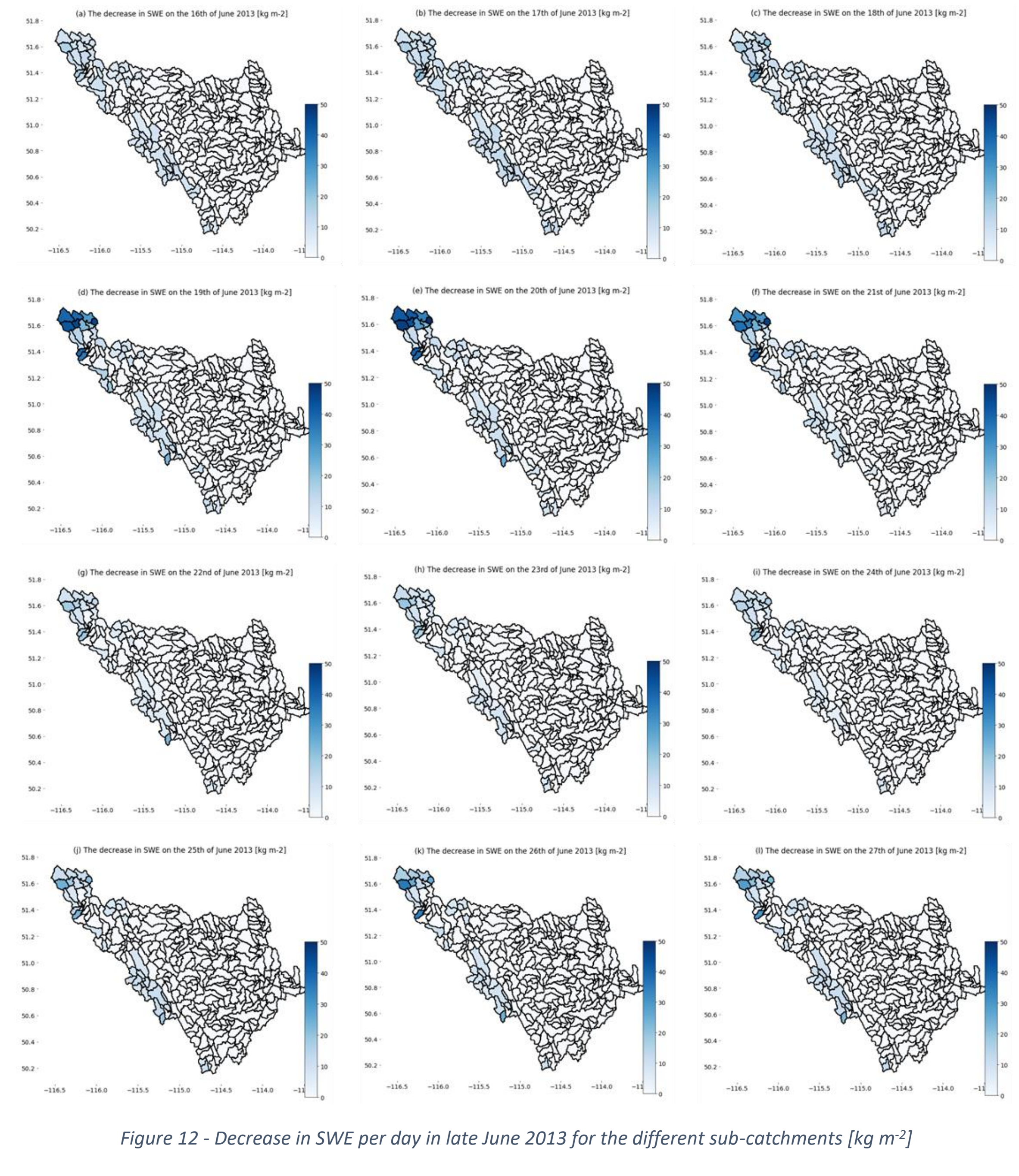

Figure 12 - Decrease in SWE per day in late June 2013 for the different sub-catchments [kg m^{-2}]

This sudden melt of the snowpack can be caused by several factors, of which the most obvious would be temperature. However, when comparing the average daily temperature in 2013 with climatology, as visualised in Figure 13a, it is not necessarily higher than average. There can be observed a lengthy increase in temperature, however, the largest part of this increase is after the flood occurrence, as can be seen in Figure 13b.

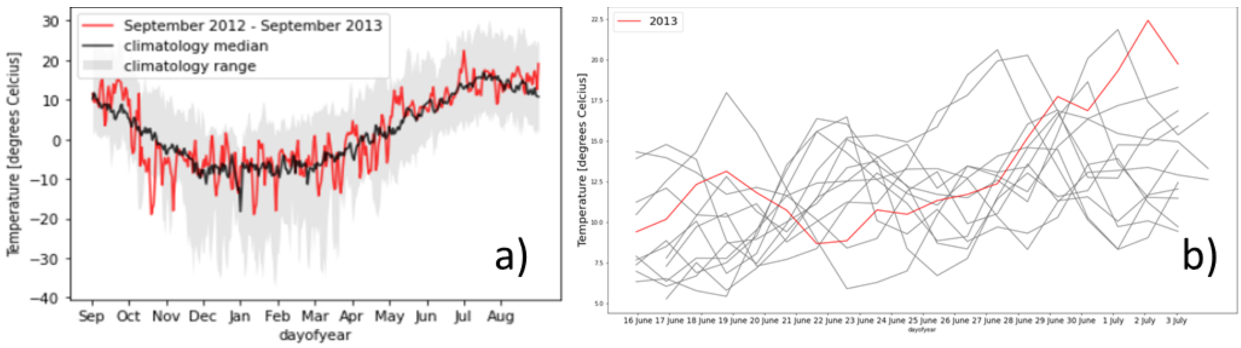

Figure 13 - Temperature in 2013 compared with climatology

When looking more in depth at the sub-catchments individually, no significant deviation from climatology can be seen as well. This suggests that the sudden increase in snow melt is caused by other factors than temperature increase. The temperature over time for the selected sub-catchments and a more detailed presentation of the selected sub-catchments with high elevation can be found in Figure 14 and Figures G-3 until G-6 in Appendix G. Note that, in Figure 14 and Figure G-3 weekly averages are used.

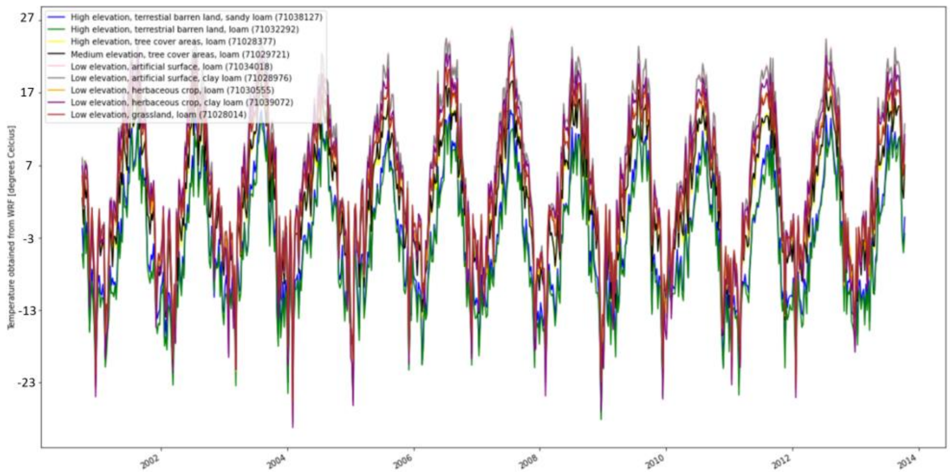

Figure 14 - Temperature over time for the selected sub-catchments

3.1.2. Precipitation

When comparing the hydrological year from September 2012 till September 2013 with previous hydrological years, starting from September 2001, the precipitation is significantly higher than average, with a large peak in late June. This is in accordance with previous studies of the 2013 Alberta flood (Pomeroy, et al., 2016) (Milrad, et al., 2015) and visualised in Figure 15.

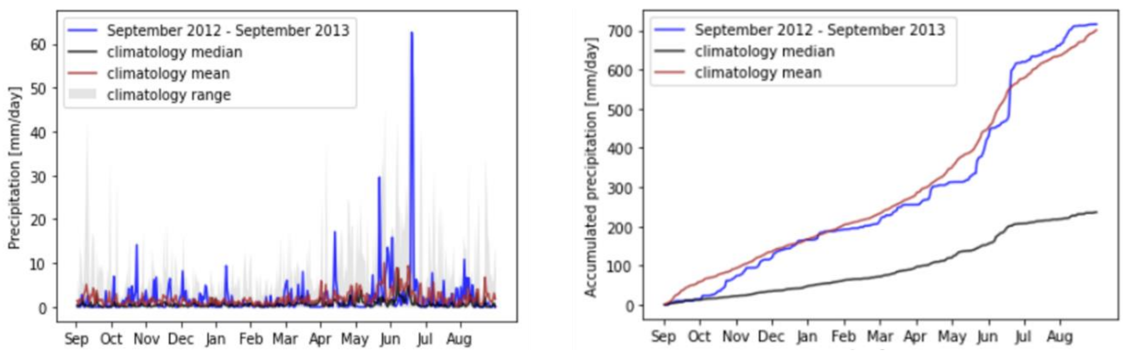

Figure 15 - Comparison of the basin average precipitation and accumulated precipitation in 2012-2013 and other hydrological years

When looking at the selected sub-catchments there are no big differences between 2013 and different years. A slight increase in precipitation can be seen in the higher elevation sub-catchments. This might be because by coincidence all the sub-catchments with increases in precipitation are not selected.

A more in depth look at the precipitation, before, during and after the flood in all the different sub-catchments shows that there was especially high precipitation in the most upstream sub-catchments of the Elbow river and the Sheep river and other medium elevation sub-catchments. Note that, the colour bar is not the same scale in each of the plots.

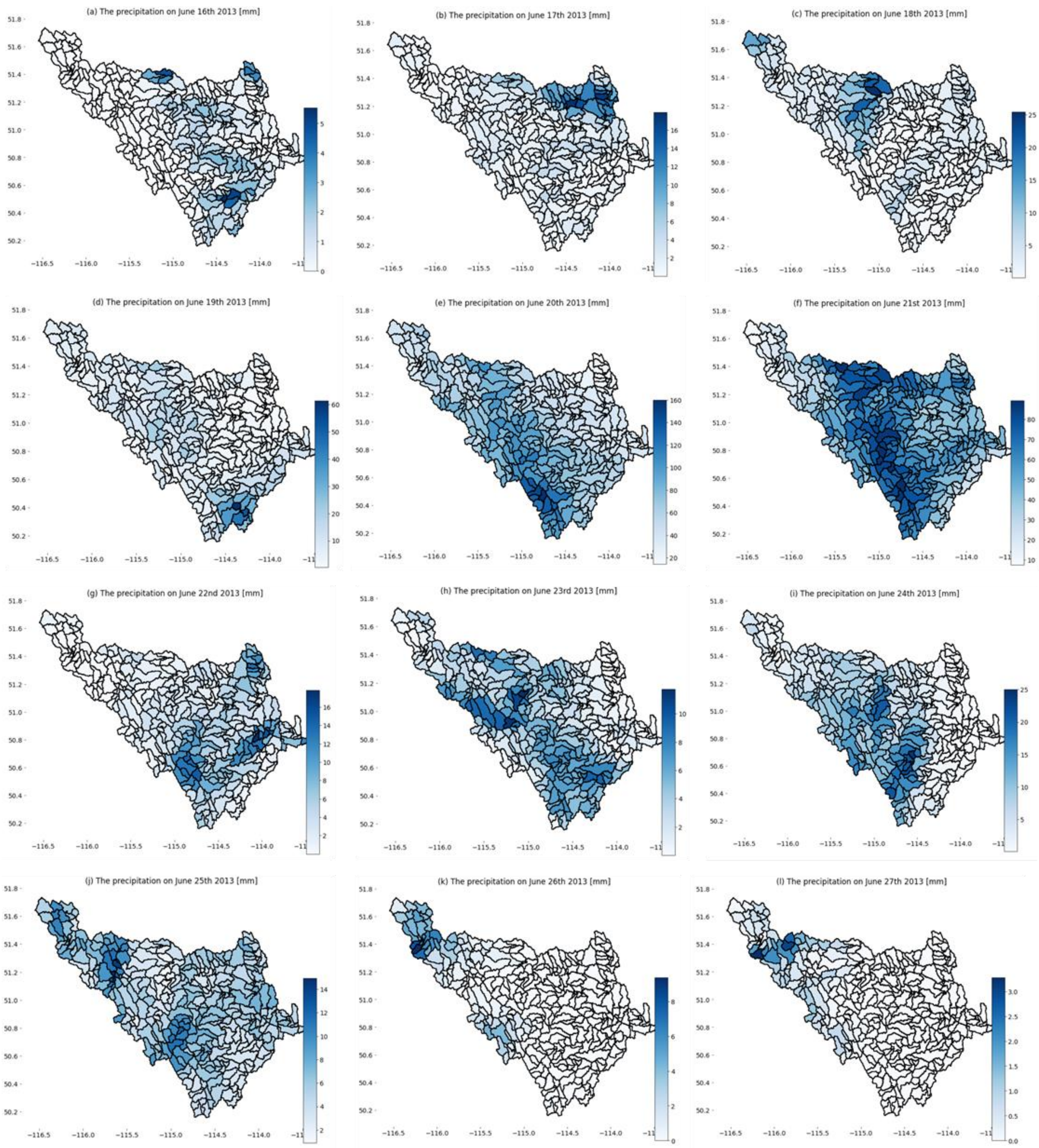

Figure 16 - Precipitation in the different sub-catchments of the study area

3.1.3. Streamflow

When looking at the simulated streamflow in the study area before, during and after the flood curve, as displayed in Figure G-7 in Appendix G, it stands out that an increased streamflow in the first side-branches of the bow river is picked up from the 20th of June, but the main increase in the discharge of the Bow river start at the 22nd of June. At first sight it seems that the precipitation has the most influence on the increase in streamflow at the outlet, but this impact also has a very steep peak. The contribution from the snow melt is overall close to the contribution of the precipitation, but more spread over days.

3.1.4. Which sub-catchments potentially contribute a lot to the flooding

From snow water equivalent analysis it shows that the days before the flooding rapid snow melt occurs in the most upstream areas of the upper Bow, therefore the river segment at Lake Louise, where the river network of those sub-catchments merges, might be an interesting segment to look at in further analysis. The Sheep river seems to contribute a lot of streamflow on the days that there is heavy precipitation. Besides that, the Bow river at Calgary and the Elbow river at Calgary are selected, before they join each other, to check the difference in the quality of flood forecasts between the river network that from simulation seems to be influenced by snow melt and that from simulation seems to be influenced by precipitation. An additional benefit of looking at the locations at Calgary is that at this location a lot of damage has been done by the 2013 Alberta Flood (Milrad, et al., 2015) (The city of Calgary, 2021). The selected sub-catchments are pictured in Figure 17.

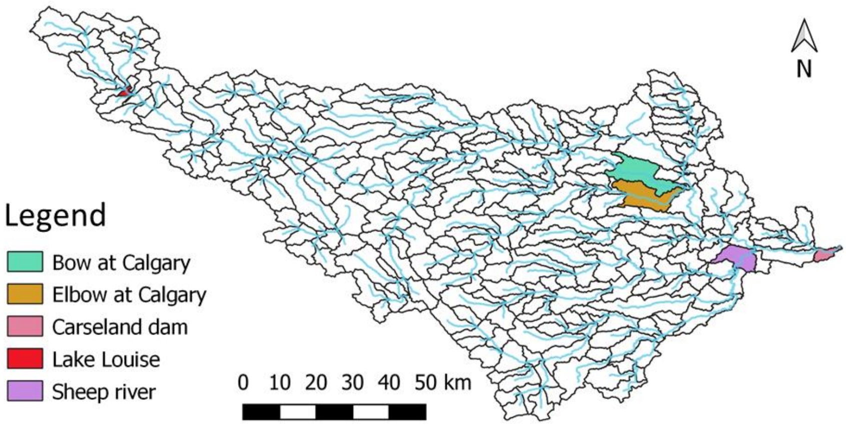

Figure 17 - Selected sub-catchments for assessment flood hindcasts

3.2. Effect of lead times on influence snow initial conditions in flood forecasts

3.2.1. Qualitative assessment of flood forecasts

When looking at the flood forecasts for each lead time for the different locations and days of the flood curve, of which a visualisation is given in Appendix F, it stands out that for almost each forecast the mean of the ensemble members predicts a discharge above the 50th percentile of the years 2000-2012 during the period 15 May till 15 July, in which seasonal floods happen in the study area (The city of Calgary, 2021). Only for the Elbow at Calgary this is not the case, for each day of the flood curve, this suggests that at this location the flood was mainly caused by precipitation, but could also be the result of the fact that in model spin-up the Elbow river is also not simulated very well for the year 2013. This means that an above average streamflow is slightly picked up by the forecasts at the other locations. However, the range of the ensemble members is still really large, till a lead time of three days, in most cases. Forecasts close to the perfect forecasts only occur with lead times that start one day before the flood curve. This could be due to the fact that a more accurate state of the river conditions results in more accurate flood forecasts (Pappenberger, et al., 2015) (Berthet, et al., 2009).

3.2.2. Continuous Ranked Probability (Skill) Scores

The overall quality of the forecasts is quantitively assessed with the continuous ranked probability skill score (CRPSS). In Figure 18 until Figure 22 the CRPSS values for each day of the flood curve are plotted. The values are depicted in Tables I-1 until I-5 in Appendix I and its corresponding values of CRPS in Tables I-6 until I-10 in Appendix I. It should be noted that, most of the locations have a smooth flood curve and include one or two days formation of the flood curve, one or two days peak of the flood curve and one or two days decline of the flood curve, however the flood curve for the Sheep river has a very steep start and both formation and the peak of the flood curve are included in the first day, where the other four days are decline of the flood curve.

When comparing the CRPSS values for the different locations it stands out that for Lake Louise the score of the CRPSS is above zero for almost every day of the flood curve and every location. The other two locations that experience the influence of the Bow river, the Bow at Calgary and the Carseland dam, also have higher values for the CRPSS, however the more downstream, the worse the scores for the CRPSS become. This suggests that better prediction of the snow melt lead to better values for the CRPSS for river segments that discharge the meltwater from this melted snow, but the more downstream, the more this influence decreases. For the Elbow river and the Sheep river the value of the CRPSS is in general close to 0 or even negative, this suggests that for those river segments the meteorological forcing almost only plays a role and thus (improved initial conditions of) the snowpack do(es) not have a large influence on the quality of the CRPSS, which is consistent with the conclusions found in the first part of this research.

For most locations and days of the flood curve, the score for the CRPSS is especially good for the lead times that start the day before the flood curve. This can again be explained by the fact that a more accurate state of the river conditions results in more accurate flood forecasts (Pappenberger, et al., 2015) (Berthet, et al., 2009). Besides that, it could be because the forecasts just did not have the time to diverge yet.

When comparing the forecasts for the different days of the flood curve, it stands out that especially for the start of the flood curve the forecasts are quite bad, but for the peak and especially the end of the flood event the forecasts are quite acceptable for some sub-catchments.

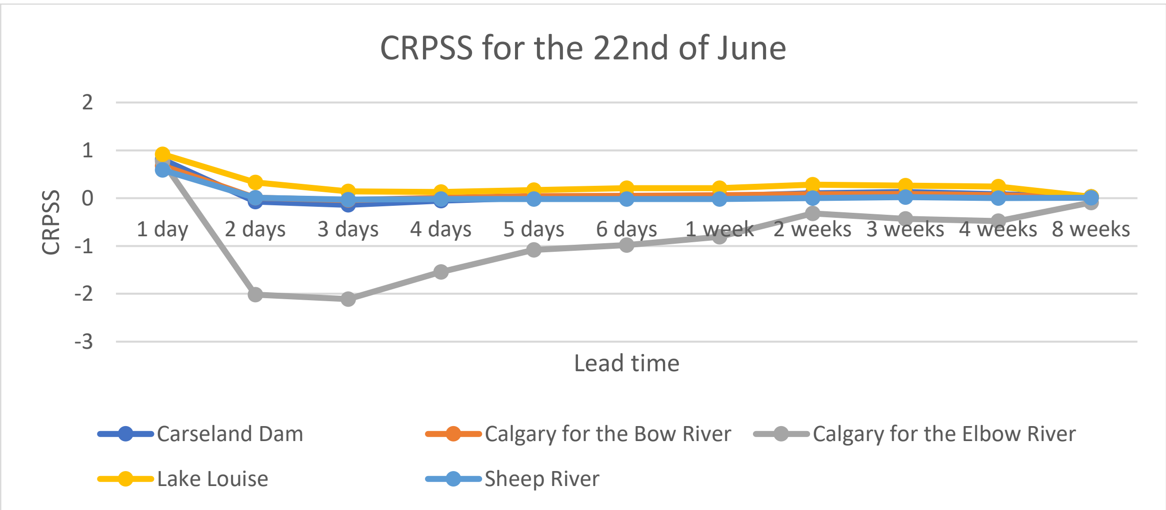

Figure 18 - CRPSS for the 22nd of June for all selected locations and lead times

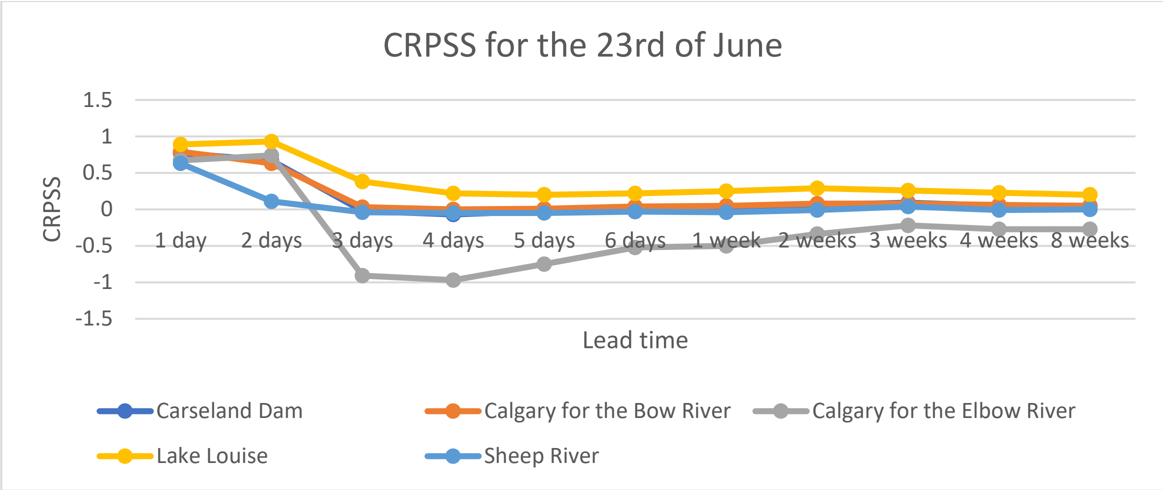

Figure 19 - CRPSS for the 23rd of June for all selected locations and lead times

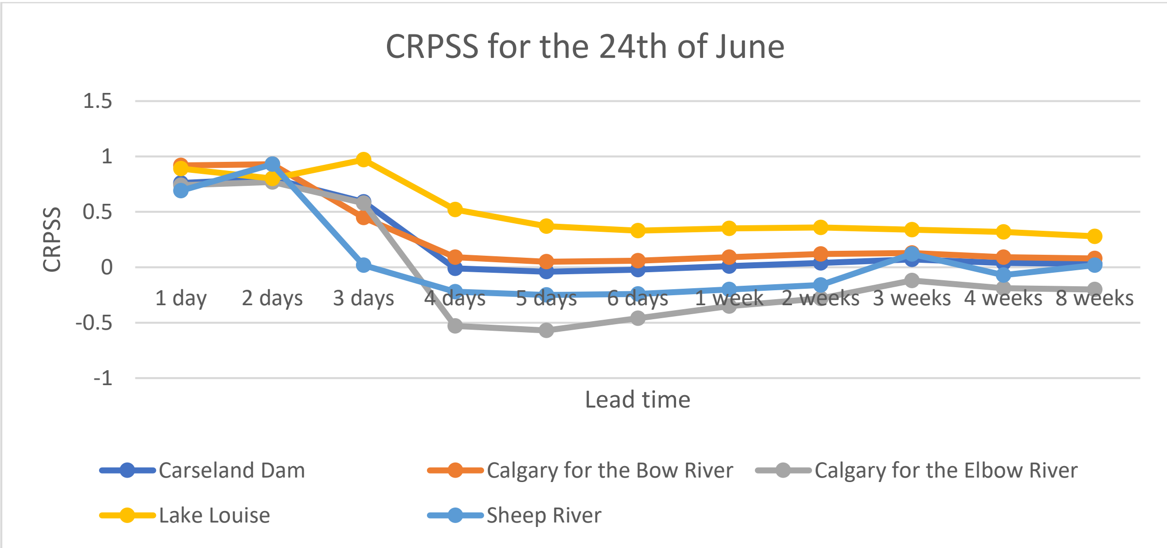

Figure 20 - CRPSS for the 24th of June for all selected locations and lead times

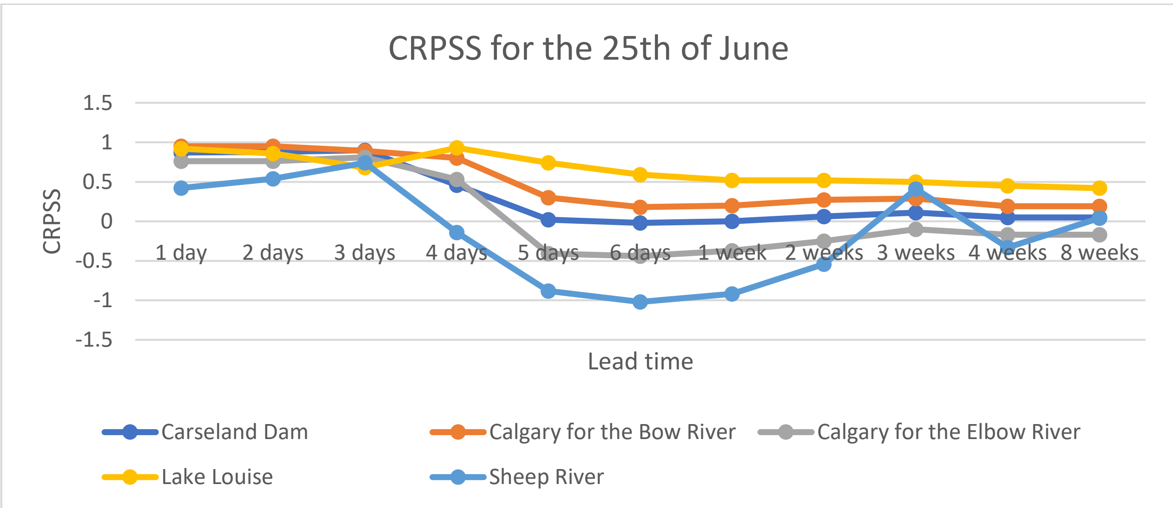

Figure 21 - CRPSS for the 25th of June for all selected locations and lead times

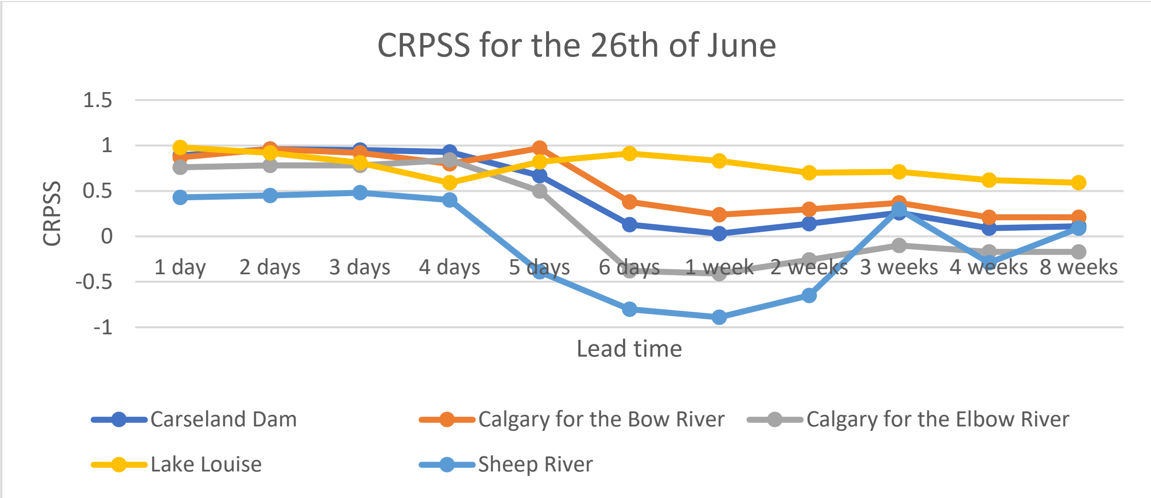

Figure 22 - CRPSS for the 26th of June for all selected locations and lead times

3.2.3. Dichotomous skill scores

In Figure 23 and Figure 24 on can find the relative operating characteristic diagram for the Bow at Calgary and the Elbow at Calgary, respectively. For clarity, the lines between the points are only plotted if they deviate from the border of the graph. The corresponding values for the plotted points are displayed in Table J-1 until Table J-10 in Appendix J. When looking at the graphs it stand out that for both the Bow and the Elbow there are almost never false alarms and the hit rate varies between 0 and 1. This means that the forecasted values for the streamflow are most of the time lower than the values of the streamflow of the 'perfect simulation'. The hit rate seems to get good values for lower thresholds and lead times that start one day before the flood. When comparing the diagrams for the Bow and the Elbow, it stands out that the resolution of the Bow seems better. This might however be caused by the fact that the streamflow of the Elbow is already largely underestimated in the 'perfect simulation'.

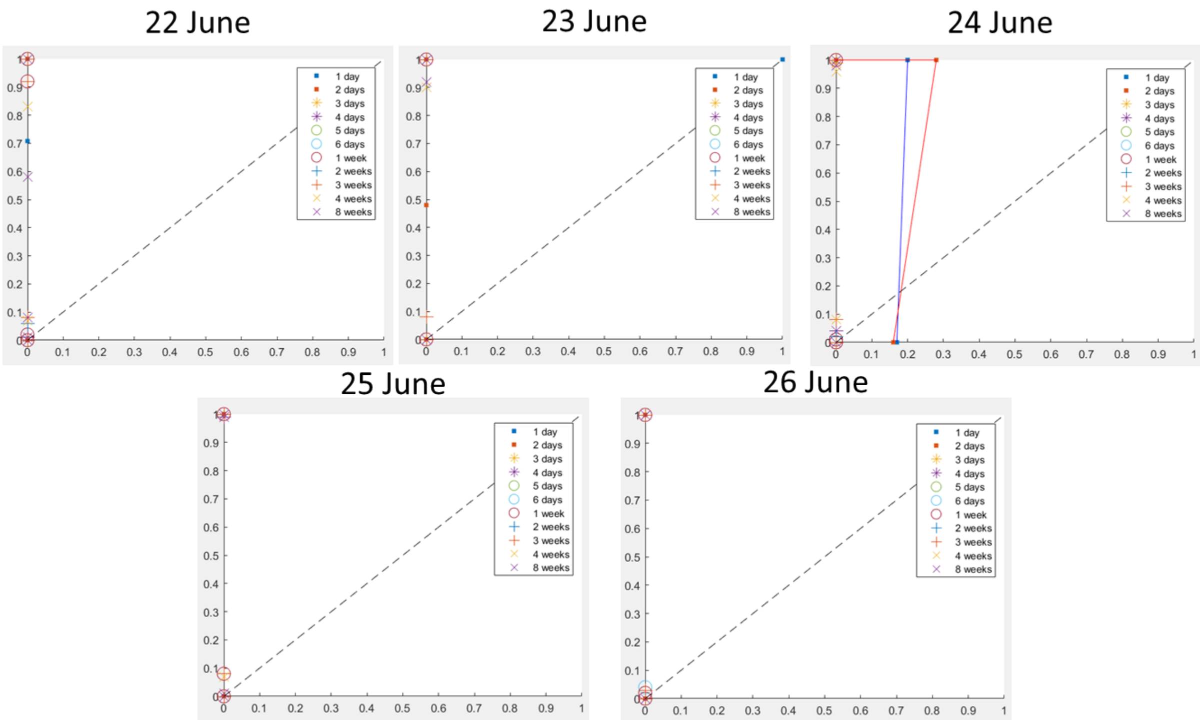

Figure 23 - ROC diagram for the Bow at Calgary (POFD on the x-axis and POD on the y-axis)

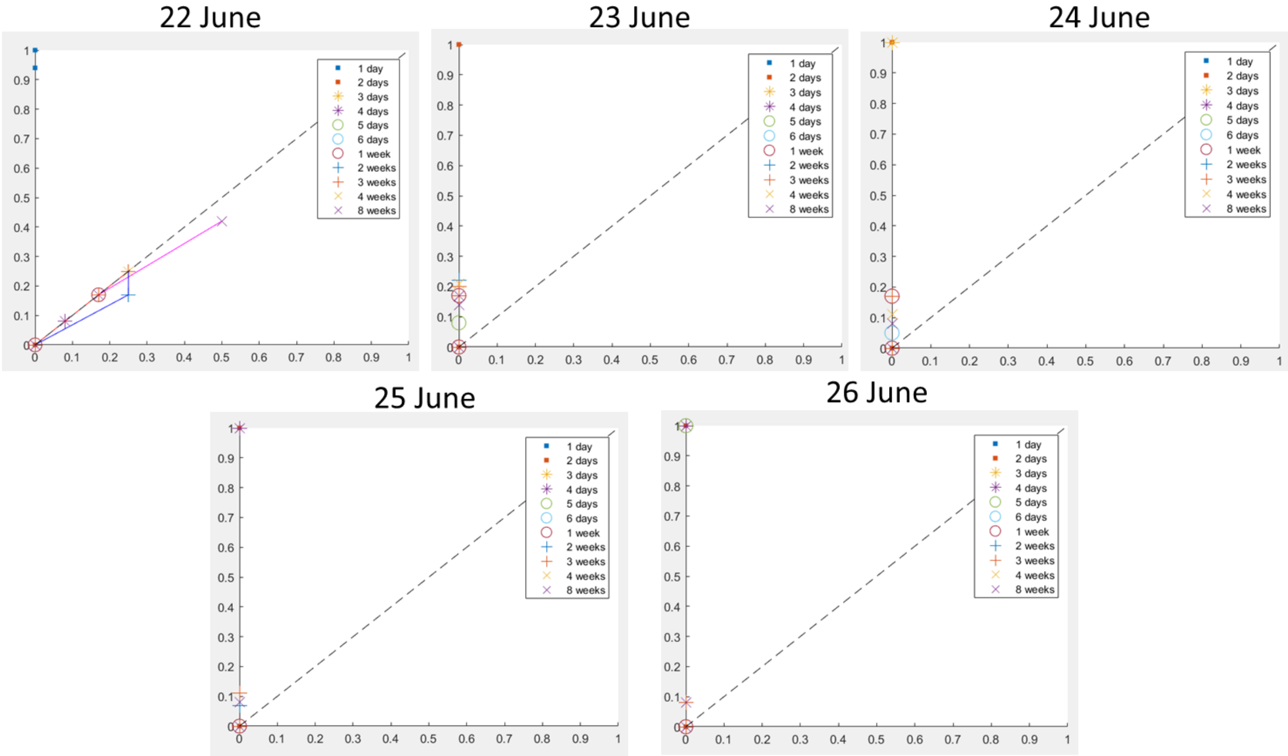

Figure 24 - ROC diagram for the Elbow at Calgary (POFD on the x-axis and POD on the y-axis)

3.2.4. Reliability diagram

In Figure 25 the reliability diagrams for the Carseland dam, the Bow at Calgary, the Elbow at Calgary, Lake Louise and the Sheep river are shown. When looking at the diagrams, the resolution is in most cases quite good. However, the sharpness and the reliability seem a bit debatable. One does however need to be careful when drawing conclusions based on those diagrams, since they contain a small number of data points (Bennett, et al., 2014).

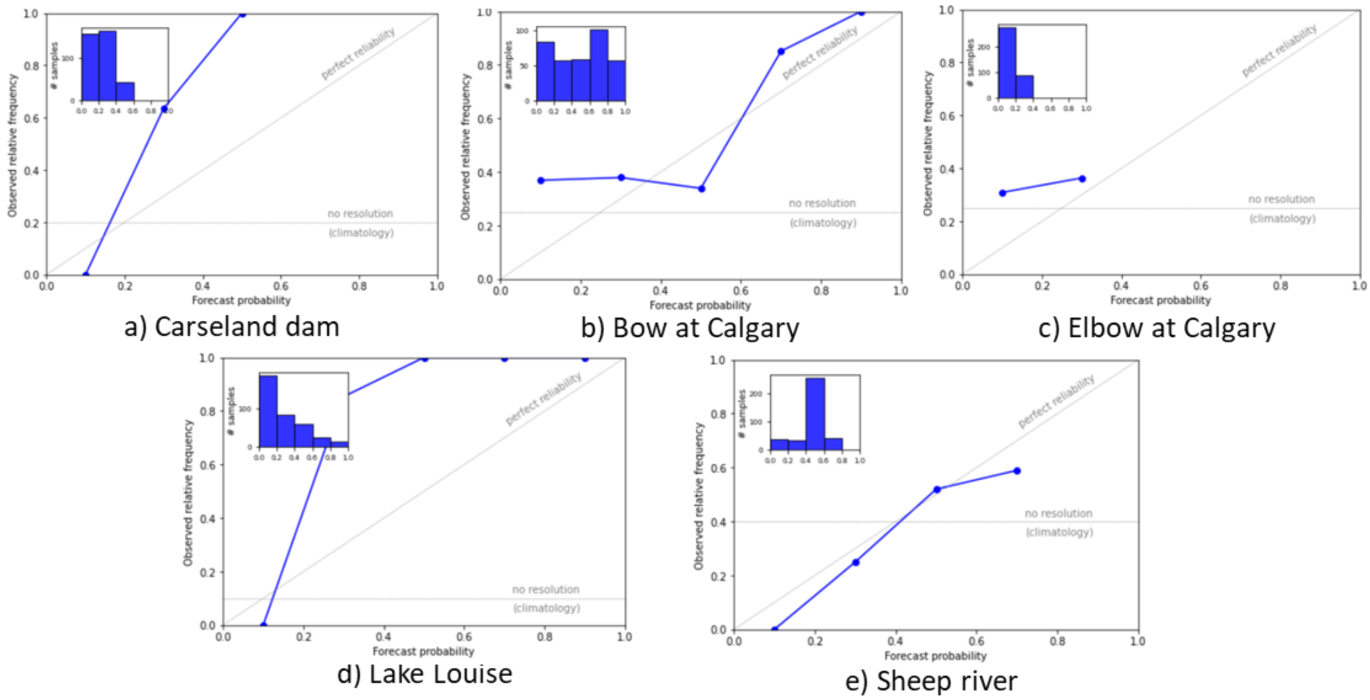

Figure 25 - Reliability diagram for all selected locations

4. Discussion

4.1. Comparison with literature

When comparing the results found in chapter 3 with literature, it stands out that both in this study and in previous studies the conditions of the snowpack seem to play a role in the unfolding of the 2013 Alberta flood event and flood forecasting for this study area (Pomeroy, et al., 2016) (Vionnet, et al., 2020). However, from the results found in this research, the influence of the snow conditions does not seem as large as described by (Pomeroy, et al., 2016). The timing of the rapid snowmelt and seems to be an important factor in the unfolding of the event, according to (Pomeroy, et al., 2016) and the results found in this research. When looking at the results found in the second part of this research, improved snow conditions might improve flood forecasts with a lead time that starts one day before the start of the flood curve. This corresponds with the statement of (Vionnet, et al., 2020), that, snow initial conditions pose a large source of uncertainty for lead times until one day. Since (Vionnet, et al., 2020) only used lead times up to 1 day, one is not able to compare longer lead times.

4.2. Limitations

4.2.1. Simplifications

A model is a simplification of reality and even though SUMMA is based on detailed representation of physical processes, it can never fully represent the real world. Physical processes are approximated with use of various equations. Not all physical processes are included in the model, glacier formation is not explicitly represented in the model and thus the representation of glaciers is not fully accurate. However, model runs show that there is a sub-catchment with eternal snow in the model. Besides that, the soil in the model is represented with an uniform depth of soil, which results in an inaccurate representation of the actual soil depth. Furthermore, SUMMA has a built in bug that it does not start simulations when canopy ice and temperature are both above 0, which seems logical, but brings problems when weather data of previous years is used for forecasting. Since ESP forecasting, the forecasting method that is used in this research, employs historic weather data, some forecasts start a day earlier or later than planned.

The study area is divided in different sub-catchments, for which the average characteristics are taken for the whole sub-catchments. Since those sub-catchments sometimes include both mountain peaks and valleys, the characteristic are not entirely correct for the whole sub-catchment. These characteristics include elevation, land-use and soil type.

4.2.2. Interpretation of results

Even though calibration is performed for the model and KGE values that in isolation can be considered acceptable (0.82, 0.72 and 0.76), the model still has a delay in the simulated streamflow and flows for the Elbow river are significantly underestimated. Therefore, conclusions drawn based on simulations and forecasts for the Elbow river, carry a lot of uncertainty. When comparing the hydrographs of this study with those used in other research, it stands out that in this research the height of the peaks is in general predicted quite good (except for the Elbow river), but the timing not that much (Yassin, 2019). Furthermore, it stands out that, in a different study they also underpredict the 2013 streamflow at the Elbow river a lot (Tesemma, et al., 2020).

The question posed in this research is what the influence of specifically snow initial conditions is on ensemble flood forecasting. However, in the execution of the research the increasing accuracy of the snow initial conditions is not isolated from increasing accuracy in other initial conditions. Literature states that increased accuracy in river state improves flood forecasts significantly. No excessive (sensitivity) analysis has been performed to prove that the improvement in flood forecasts for shorter lead times are caused by only the increased accuracy of the snow initial conditions.

4.2.3. Forcing data

This study uses historic weather data for forecasting, which might be less accurate than (an ensemble of) real weather forecasts. For future research one might want to use real weather forecasts. This might bring some challenges. For the simulations a lot of input parameters were required, which are not always easily accessible. For hindcasting with historic data, the WRF data was suitable, since it contains all the required data and is available for the whole Northern American continent. However, for further research acquiring weather data that contains all these variables might be challenging, especially for longer lead times.

5. Conclusion and recommendations

5.1. Conclusion

When comparing the hydrological year from September 2012 till September 2013 with previous hydrological years, starting from September 2001, it seems that the snow water equivalent is not necessary higher than previous years, but that the snow melt starts earlier in the year and goes more rapidly. In the same days as the rapid snow melt, heavy precipitation occurs in the front ranges of the study area. This suggests that the important factor in the unfolding of the 2013 flood event was unlucky timing of several events. However, the rapid snow melt was not caused by a sudden increase in temperature, illustrating the influence of a uncertain snowpack.

When looking at the flood forecasts for different lead times and different days of the flood peak, it stands out that the flood forecasts only scores well when the forecast is initialised one day before the simulated flood starts. This could mean that the flood forecasts only becomes better because of the improved accuracy in the conditions of the flow in the river channel. However, a slight differentiation in quality of the flood forecasts can be seen in the river segments that contain the streamflow that resulted from the rapid snowmelt and river segments that contain the streamflow that resulted from heavy precipitation. This suggests that the conditions of the snowpack do have influence on ensemble flood forecasting, however, little.

5.2. Recommendations

Further research needs to be done to determine if improved river network states or improved snowpack conditions contributed the most to improved flood forecasts. Because of that, in this section only recommendations are done for future research. When those points are researched, one can give recommendations on how to improve snow measurements and operational forecasting systems in Alberta.

The research performed as described has a couple of limitations, of which some lead to recommendations for further research. The most important is that in this research there is only looked into the potential influence of the snow initial conditions. When one wants to go more in depth on the influence of the snow initial conditions it is advised to use different sets of initial conditions for each starting date, instead of the simulated initial conditions from 2013. Furthermore, it is recommended to investigate what conclusions can be found when performing the steps in this research with a real weather model for its ensemble members, instead of meteorological data from previous years. Besides that, it is advised to look into how spatial discretization of snow initial conditions affects ensemble flood forecasting. In this study the model is divided into different sub-catchments, the study area could be further discretised based on elevation bands.

References

Allen, G. (2021, March 28). *MERIT Hydro Visualization and Interactive Map*. Retrieved from MERIT Hydro: global hydrography datasets: https://meritdataset.users.earthengine.app/view/merit-hydro-visualization-and-interactive-map

Bash, E. A., & Marshall, S. J. (2014). Estimation of glacial melt contributions to the Bow River, Alberta, Canada, using a radiation-temperature melt model. *Annals of glaciology*, 138-152.

Bennett, A., Wayand, N., Nijssen, B., Clark, M., & Nearing, G. (2021, October 19). *summa*. Retrieved from GitHub: https://github.com/CH-Earth/summa/blob/master/build/source/dshare/var_lookup.f90

Bennett, J. C., Robertson, D. E., Shrestha, D. L., Wang, Q. J., Enever, D., Hapuarachchi, P., & Tuteja, N. K. (2014). A System for Continuous Hydrological Ensemble Forecasting (SCHEF) to lead times of 9 days. *Journal of Hydrology*, 2832–2846.

Berthet, L., Andreassian, V., Perrin, C., & Javelle, P. (2009). How crucial is it to account for the antecedent moisture conditions in flood forecasting? Comparison of event-based and continuous approaches on 178 catchments. *Hydrology and Earth System Sciences*, 819-831. Retrieved from https://www.sciencedirect.com/science/article/pii/S0022169415000414#b0020

Brier, G. (1950). Verification of forecasts expressed in terms of probability. *Monthly Weater Review*(78). Retrieved from https://books.google.nl/books?hl=nl&lr=&id=jnbpAAAAMAAJ&oi=fnd&pg=RA1-PA1&ots=0X--1JpNwN&sig=caQ0kABNZPfWT3Wr76P_mFnf8d4&redir_esc=y#v=onepage&q&f=false

Clark, M. P. (2021, October 19). *Initial conditions, restart or state file*. Retrieved from summa: https://summa.readthedocs.io/en/latest/input_output/SUMMA_input/#initial-conditions-restart-or-state-file

Clark, M. P., Nijssen, B., Lundquist, J., Kavetski, D., Rupp, D., Woods, R., . . . Rasmussen, R. (2015a). A unified approach to process-based hydrologic modeling. Part 1: Modeling concept. *Water Resources Research*(51). doi:10.1002/2015WR017198

Clark, M., Nijssen, B., Lundquist, J. D., Kavetski, D., Rupp, D. E., Woods, R. A., . . . Marks, D. G. (2015c). *The structure for unifying multiple modeling alternatives (SUMMA), Version 1.0: Technical Description.* NCAR Technical Note NCAR/TN-514+STR. doi:10.5065/D6WQ01TD

Clark, M., Nijssen, B., Lundquist, J., Kavetski, D., Rupp, D., Woods, R., . . . Marks, D. (2015b). A unified approach for process-based hydrologic modeling: 2. Model implementation and case studies. *Water Resources Research*(51). doi:10.1002/2015WR017200

Cloke, H., & Pappenberger, F. (2009). Ensemble flood forecasting: A review. *Journal of Hydrology*, 613-626.

Day, G. (1985). Extended streamflow forecasting using NWSRFS. *Journal of Water Resources Planning and Management*(111), 157-170. doi:10.1061/(ASCE)0733-9496(1985)111:2(157)

Dornes, P. F., Pomeroy, J. W., Pietroniro, A., & Verseghy, D. L. (2008). *Effects of Spatial Aggregation of Initial Conditions and Forcing Data on Modeling.* Saskatoon: American Meteorological Society. doi: 10.1175/2007JHM958.1

ESA. (2021, May 25). *ESA/CCI viewer*. Retrieved from http://maps.elie.ucl.ac.be/CCI/viewer/download.php

F.Pappenberger, M.H.Ramos, Cloked, H., F.Wetterhall, Alfieri, L., K.Bogner, . . . P.Salamon. (2015). How do I know if my forecasts are better? Using benchmarks in hydrological ensemble prediction. *Journal of hydrology*(522), 697-713. doi:https://doi.org/10.1016/j.jhydrol.2015.01.024

Free Map Tools. (2021, June 7). *Find Elevation Map*. Retrieved from https://www.freemaptools.com/elevation-finder.htm

Google. (2021, March 29). Retrieved from Google Earth: https://earth.google.com/web/@51.36831594,-115.11829697,1696.60079926a,271444.77011729d,30.00000274y,0.0000246h,0t,0r

Government of Canada. (2021, April 20). *Canadian historical snow survey data*. Retrieved from Environment and Climate Change Canada Data: https://data.ec.gc.ca/data/climate/systems/canadian-historical-snow-survey-data/?wbdisable=true

Government of Canada. (2021, April 26). *Canadian historical snow survey data*. Retrieved from Environment and Climate Change Canada Data: https://data.ec.gc.ca/data/climate/systems/canadian-historical-snow-survey-data/?wbdisable=true

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. 80-91. doi:10.1016/j.jhydrol.2009.08.003

Hersbach, H. (2000). Decomposition of the Continuous Ranked Probability Score for Ensemble Prediction Systems. *Weather and Forecasting, 15*(5), 559–570. doi:https://doi.org/10.1175/1520-0434(2000)015<0559:DOTCRP>2.0.CO;2

Knoben, W. J., Freer, J. E., & Woods, R. A. (2019). Technical note: Inherent benchmark or not? Comparing NashSutcliffe and Kling-Gupta efficiency scores. *Hydrology and earth systems sciences*. doi:https://doi.org/10.5194/hess-2019-327

Krause, P., Boyle, D. P., & Bäse, F. (2005). Comparison of different efficiency criteria for hydrological model assessment . *Advances in geosciences, European geosciences Union*, 89-97. Retrieved from https://hal.archives-ouvertes.fr/hal-00296842

Li, H., Luo, L., Wood, E. F., & Schaake, J. (2009). The role of initial conditions and forcing uncertainties in seasonal hydrologic forecasting. *Journal of geophysical research*. doi:doi:10.1029/2008JD010969

Lin, P., Pan, M., Beck, H. E., Yang, Y., Yamazaki, D., Frasson, R., . . . Wood, E. F. (2019). Global Reconstruction of Naturalized River Flows at 2.94 Million Reaches. *Water resources research*, 6499-6516.

Lopez, M. G., Crochemore, L., & Pechlivanidis, I. G. (2021). Benchmarking an operational hydrological model for providing seasonal forecasts in Sweden. *Hydrology and earth systems sciences*(25), 1189-1209. doi:https://doi.org/10.5194/hess-25-1189-2021

McClave, J. T., & Sincich, T. (2013). *Statistics.* Pearson Education, Limited.

Milrad, S. M., Gyakum, J. R., & Atallah, E. H. (2015). A Meteorological Analysis of the 2013 Alberta Flood: Antecedent Large-Scale Flow Pattern and Synoptic–Dynamic Characteristics. *Montly Weather Review*, 2817-2841. doi:https://doi.org/10.1175/MWR-D-14-00236.1

Mizukami, N., Clark, M. P., Sampson, K., Nijssen, B., Mao, Y., McMillan, H., . . . Brekke, L. D. (2016). mizuRoute version 1: a river network routing tool for a continental domain water resources applications. *Geosci. Model Dev.*(9), 2223–2238.

National Resources Canada. (2021, April 12). *Rivers*. Retrieved from The Atlas of Canada: https://web.archive.org/web/20070404150649/http://atlas.nrcan.gc.ca/site/english/learningresources/facts/rivers.html

Pomeroy, J. W., Stewart, R. E., & Whitfield, H. P. (2016). The 2013 flood event in the South Saskatchewan and Elk River basins: Causes, assessment and damages. *Canadian Water Resources Journal*, 105-117. doi:https://doi.org/10.1080/07011784.2015.1089190

Powers, J. G., Klemp, J. B., Skamarock, W. C., Davis, C. A., Dudhia, J., Gill, D. O., . . . Duda, M. G. (2017). The Weather Research and Forecasting Model: Overview, System Efforts, and Future Directions. *Bulletin of the American Meteorological Society*, 1717-1737. doi:https://doi.org/10.1175/BAMS-D-15-00308.1

PyPI. (2021, May 22). *Proper scoring rules in Python.* Retrieved from https://pypi.org/project/properscoring/

Rakovec, O. (2014). *Improving operational flood forecasting using data assimilation.* Wageningen University. Retrieved from https://edepot.wur.nl/298593

Ranjan, R. (2009). *Combining and Evaluating Probabilistic Forecasts.* University of Washington.

Rasmussen, R., & Liu, C. (2021, April 23). *High Resolution WRF Simulations of the Current and Future Climate of North America*. doi:https://doi.org/10.5065/D6V40SXP

Shaw, E., Beven, K., Chappel, N., & Lamb, R. (2011). *Hydrology in practice.* New York: Spon press.

Tesemma, Z., Shook, K., Princz, D., Razavi, S., Wheater, H., Davison, B., . . . Pomeroy, J. (2020). *Diagnosis of Historical and Future Flow Regimes of the Bow River at Calgary Using a Dynamically Downscaled Climate Model and a Physically Based Land Surface Hydrological Model.* Saskatoon: University of Saskatchewan.

The city of Calgary. (2021, June 20). *Flood Readiness: Understand. Prepare. Stay Informed, Bow and Elbow Rivers, Flows, Triggers and Related Effects.* Retrieved from https://www.calgary.ca/uep/water/flood-info/types-of-flooding-in-calgary/understanding-river-flow-rates.html

The city of Calgary. (2021, June 2). *Why does flooding happen in Calgary?* Retrieved from Calgary: https://www.calgary.ca/uep/water/flood-info/types-of-flooding-in-calgary/understanding-river-flow-rates.html

Tolson, B. A., & Shoemaker, C. A. (2007). Dynamically dimensioned search algorithm for computationally efficient watershed model calibration. *Water Resources Research*(43). doi:https://doi.org/10.1029/2005WR004723

University of Saskatchewan. (2021, April 9). *Centre for Hydrology, University of Saskatchewan*. Retrieved from Centre for Hydrology: https://research-groups.usask.ca/hydrology/index.php

University of Saskatchewan. (2021, April 9). *Welcome to Computational Hydrology within the Centre for Hydrology at the University of Saskatchewan*. Retrieved from Computational Hydrology: https://uofs-comphyd.github.io/

Vionnet, V., Fortin, V., Gaborit, E., Roy, G., Abrahamowicz, M., Gasset, N., & Pomeroy, J. W. (2020). Assessing the factors governing the ability to predict late-spring flooding in cold-region mountain basins. *Hydrology and Earth System Sciences*(24), 2141-2165. doi:https://doi.org/10.5194/hess-24-2141-2020

Wood, A. W., Hopson, T., Newman, A., Brekke, L., Arnold, J., & Clark, M. (2016). Quantifying Streamflow Forecast Skill Elasticity to Initial Condition and Climate Prediction Skill. *Journal of Hydrometerology*, 651-668. doi:https://doi.org/10.1175/JHM-D-14-0213.1

World Meteorological Organization. (2012). *Guidelines on Ensemble Prediction Systems and Forecasting.* Geneva. Retrieved from https://library.wmo.int/index.php?lvl=notice_display&id=12962#.YW6EEBpBxPY

World Weather Research Program. (2021, March 8). *Standard verification methods*. Retrieved from WWRP/WGNE Joint Working Group on Forecast Verification Research: https://www.cawcr.gov.au/projects/verification/#Eyeball_method

Wu, W., Emerton, R., Duan, Q., Wood, A. W., Wetterhall, F., & Robertson, D. E. (2020). *Ensemble flood forecasting: Current status and future opportunities.* Wiley. doi:10.1002/wat2.1432

Yamazaki, D., Ikeshima, D., Sosa, J., Bates, P. D., Allen, G. H., & Pavelsky, T. M. (2019). MERIT Hydro: A High-Resolution Global Hydrography Map Based on Latest Topography Dataset. *Water resources research*, 5053-5073.

Yassin, F. (2019). *TOWARDS IMPROVED HYDROLOGIC LAND SURFACE MODELLING: ENHANCED MODEL IDENTIFICATION AND INTEGRATION OF WATER MANAGEMENT.* Saskatoon: University of Saskatchewan.

Zahmatkesh, Z., Jha, S., Paulin, C., & Stadnyk, T. (2019). An overview of river flood forecasting procedures. *Canadian Water Resources Journal*. doi:https://doi.org/10.1080/07011784.2019.1601598

Appendix A– Technical details SUMMA and mizuRoute

General technical details SUMMA

When using SUMMA there are a couple of simulation files that the model needs to run its simulation. Those files are managed with the fileManager, which contains the path to the different input files and specifies the dates at which the simulation needs to be run. Since SUMMA supports different modelling options, the modelling options that need to be used by SUMMA should be specified in a text file, to which is directed by the fileManager. Furthermore a list of all the forcing files that SUMMA should use needs to be included as a text file. The output that SUMMA gives is dependent on the output control file, in which you can specify which variables should be saved during the simulations (Clark, et al., 2015c).

The simulation can be started either with or without setting initial conditions. When those initial conditions are set the model has no warm-up period, which is especially important for flood forecasting. When running a simulation (with or without setting initial conditions) a restart file can be saved, at the end of the simulation, daily, monthly or yearly. This restart file can then be directed to, to represent the initial conditions for a new simulation (Clark, et al., 2015c).

Creating the ensemble members SUMMA setting files for each of the simulations

For each of the days of the flood and for each lead time, new setting files needed to be created, since the combination of start and end date for each of those simulations is different. The WRF reanalysis, as earlier described in section 2.1.3., will be used as meteorological forcing data for the flood forecasts. Each from the years of the reanalysis, 2001-2013, will be used as an individual ensemble member. This is done by creating a separate fileManager text file for each year.

General technical details mizuRoute

For mizuRoute to work properly there are a couple of files needed. To run the routing simulations mizuRoute needs a file with the network topology, a routing control file and a parameter file. The routing control file includes; the start and end date of the simulation, route options, file paths, runoff data and output options. The parameter file can either contain default values or values obtained from an earlier simulation. The routing conditions can be saved at the end of the simulation, yearly, monthly or daily. The run-off for a certain sub-catchment can be determined by accessing the IRFroutedRunoff of the concerning sub-catchment within mizuRoute.

Creating the ensemble members mizuRoute setting files for each of the simulations

For each of the thirteen ensemble members the routing between the different sub-catchments needs to be assigned. This is done by creating routing control files for each of the days of the flood and for each lead time, since the combination of start and end date for each of those simulations is different.

Appendix B - Model set-up

Table B-1 - Modelling decisions SUMMA (Clark, et al., 2015c)

Modelling choice	**Explanation**	**Selected option**
soilCatTbl	soil-category dataset	ROSETTA
vegeParTbl	vegetation category dataset	MODIFIED_IGBP_MODIS_NOAH
soilStress	choice of function for the soil moisture control on stomatal resistance	NoahType
stomResist	choice of function for stomatal resistance	BallBerry
num_method	choice of numerical method	iterative
fDerivMeth	method used to calculate flux derivatives	analytic
LAI_method	method used to determine LAI and SAI	monTable
f_Richards	form of Richard's equation	mixdform
groundwatr	choice of groundwater parameterization	bigBuckt
hc_profile	choice of hydraulic conductivity profile	constant
bcUpprTdyn	type of upper boundary condition for thermodynamics	ngr_flux
bcLowrTdyn	type of lower boundary condition for thermodynamics	zeroFlux
bcUpprSoiH	type of upper boundary condition for soil hydrology	liq_flux
bcLowrSoiH	type of lower boundary condition for soil hydrology	Drainage
veg_traits	choice of parameterization for vegetation roughness length and displacement	CM_QJRMS1988
canopyEmis	choice of parameterization for canopy emissivity	difTrans
snowIncept	choice of parameterization for snow interception	lightSnow
windPrfile	choice of wind profile through the canopy	logBelowCanopy
astability	choice of stability function	Louisinv
canopySrad	choice of canopy shortwave radiation method	CLM_2stream
alb_method	choice of albedo representation	conDecay
compaction	choice of compaction routine	Anderson
snowLayers	choice of method to combine and sub-divide snow layers	CLM_2010
thCondSnow	choice of thermal conductivity representation for snow	jrdn1991
thCondSoil	choice of thermal conductivity representation for soil	mixConstit

spatial_gw	choice of method for the spatial representation of groundwater	localColumn
subRouting	choice of method for sub-grid routing	timeDlay

Table B-2 - Overview of variables in WRF dataset (Clark M. , et al., 2015c)

Variable	**Explanation**	**Dependent on**
Latitude	-	Hydrological response unit
Longitude	-	Hydrological response unit
LWRadAtm	Longwave radiation	Hydrological response unit & time
airpres	Air pressure	Hydrological response unit & time
pptrate	Precipitation rate	Hydrological response unit & time
spechum	Specific humidity	Hydrological response unit & time
SWRadAtm	Shortwave radiation	Hydrological response unit & time
airtemp	Air temperature	Hydrological response unit & time
windspd	Wind speed	Hydrological response unit & time

Table B-3 - Parameters used for calibration (Bennett, et al., 2021)

Parameters	**Explanation**
k_macropore	Saturated hydraulic conductivity for macropores (m s-1)
k_soil	Hydraulic conductivity of soil (m s-1)
theta_sat	Soil porosity (-)
aquiferBaseflowExp	Baseflow rate when aquifer storage = aquiferScaleFactor (m s-1)
aquiferBaseflowRate	Baseflow exponent (-)
qSurfScale	Scaling factor in the surface runoff parameterization (-)
summerLAI	Maximum leaf area index at the peak of the growing season (m2 m-2)
frozenPrecipMultip	Frozen precipitation multiplier (-)
heightCanopyBottom	Height of bottom of the vegetation canopy above ground surface (m)
thickness (heightCanopyTop - heightCanopyBottom)	Effectively calibrates parameter heightCanopyTop, but in this way it ensures that canopy top > canopy bottom. heightCanopyTop: height of top of the vegetation canopy above ground surface (m)
routingGammaScale	Scale parameter in Gamma distribution used for sub-grid routing (s)
routingGammaShape	Shape parameter in Gamma distribution used for sub-grid routing (-)

Appendix C – Model and input validation

Model validation

Snow water equivalent (SWE)

Table C-1 - Characteristics snow observation stations

Station ID	**Station name**	**Elevation of station [m]** (Free Map Tools, 2021)	**Corresponding HRU**	**HRU mean elevation [m]** (Allen, 2021) (Yamazaki, et al., 2019)
ALE-05BF824P	THREE ISLE LAKE PILLOW	2165	71038383	2324
ALE-05BB803P	SUNSHINE VILLAGE PILLOW	2209	71030690	2163
ALE-05CA805P	SKOKI LODGE PILLOW	2089	71032102	2302
ALE-05BJ805P	LITTLE ELBOW SUMMIT PILLOW	2159	71038629	2285
ALE-05BL811P	LOST CREEK SOUTH PILLOW	2108	71037075	2037
ALE-05BL812P	MOUNT ODLUM PILLOW	2079	71036170	2059

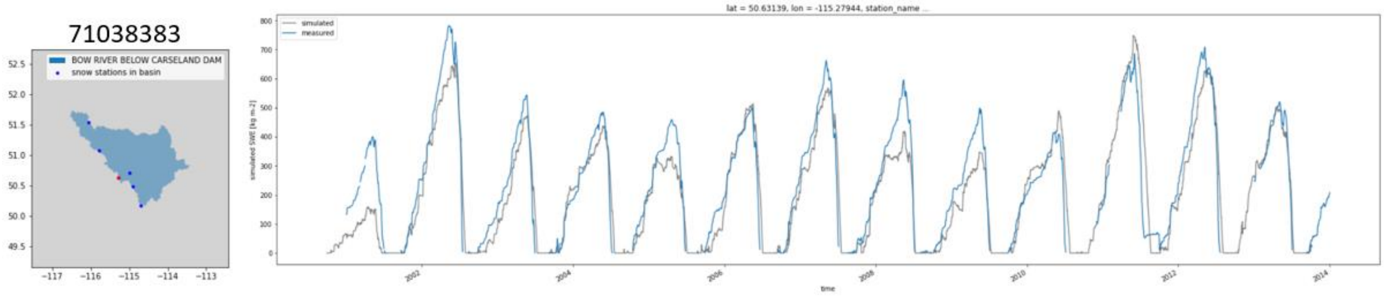

Figure C-1 - Difference between SWE observation and simulation for Three Isle Lake Pillow

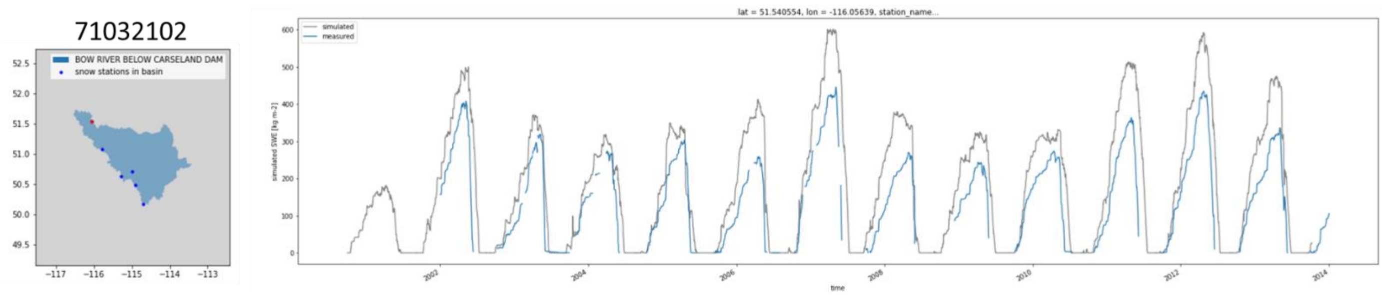

Figure C-2 - Difference between SWE observation and simulation for Skoki Lodge Pillow

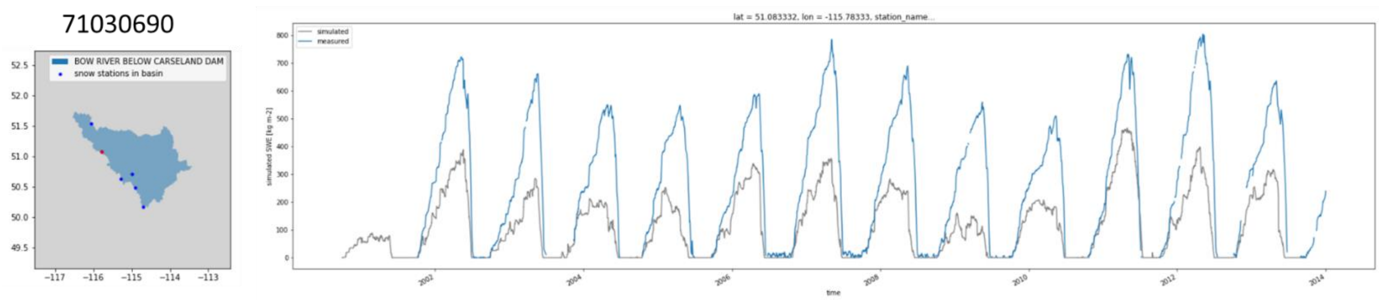

Figure C-3 - Difference between SWE observation and simulation for Sunshine Village Pillow

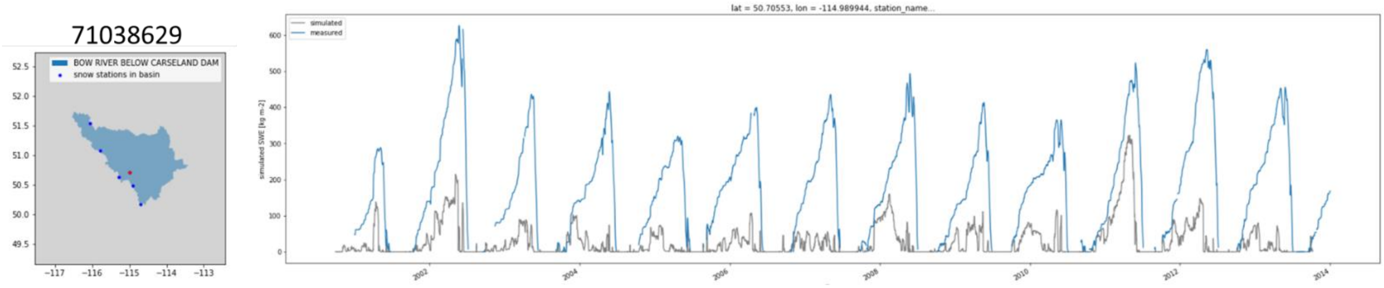

Figure C-4 - Difference between SWE observation and simulation for Little Elbow Summit Pillow

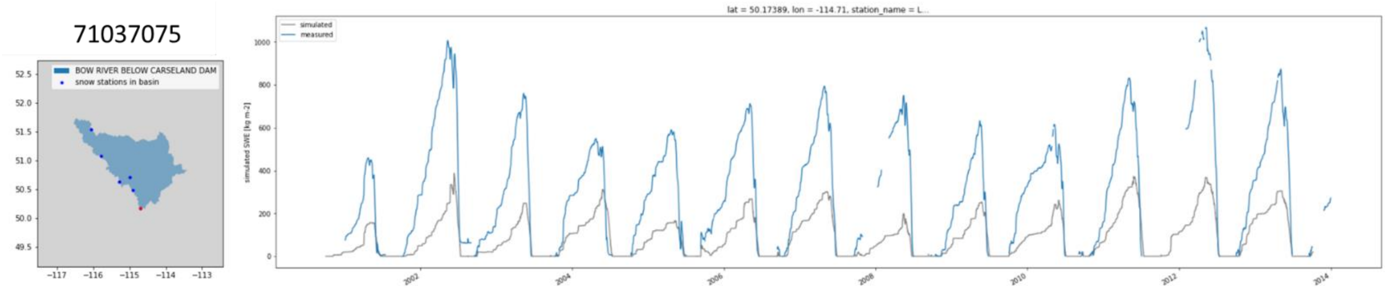

Figure C-5 - Difference between SWE observation and simulation for Lost Creek South Pillow

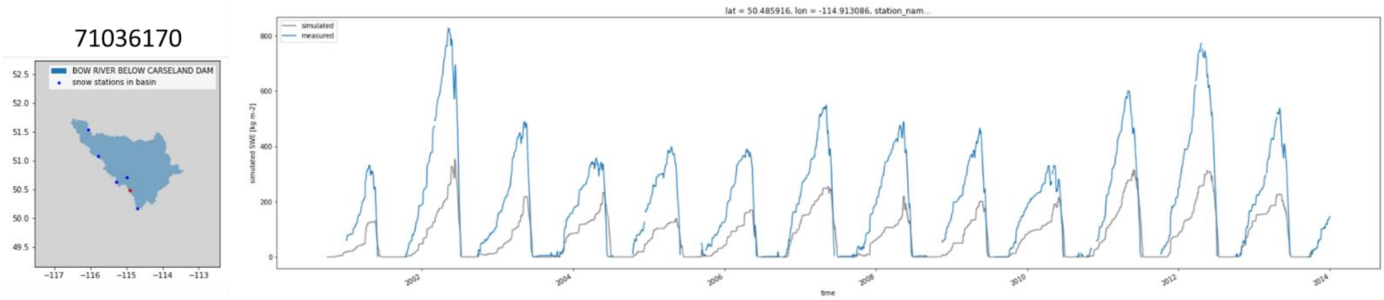

Figure C-6 - Difference between SWE observation and simulation for Mount Odlum Pillow

Streamflow

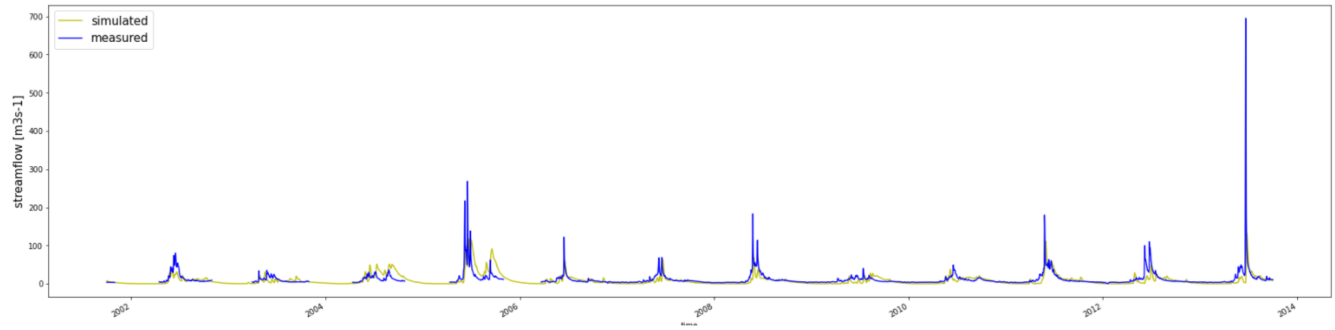

Figure C-7 - Difference between the simulated and measured streamflow for the Elbow 2001-2013

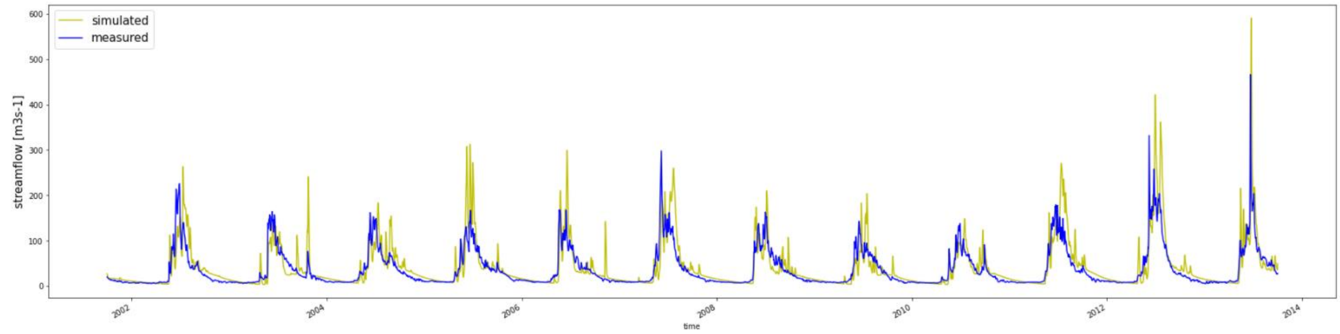

Figure C-8 - Difference between the simulated and measured streamflow for the Bow at Banff 2001-2013

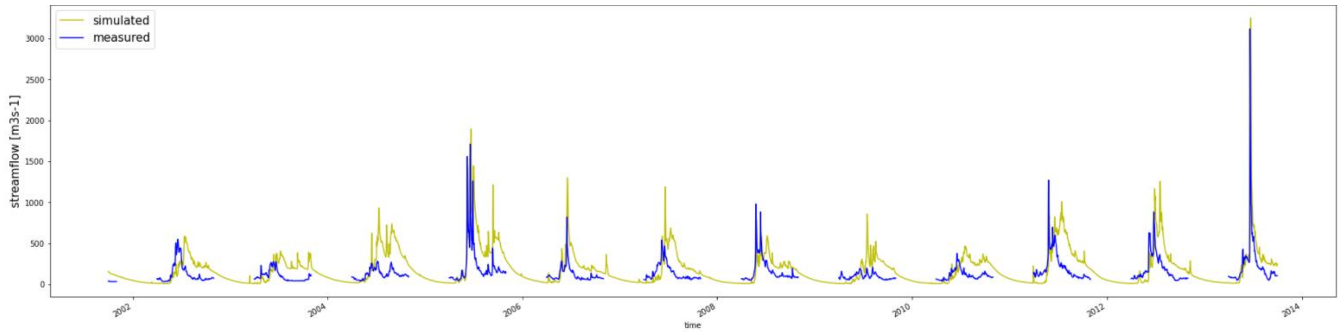

Figure C-9 - Difference between the simulated and measured streamflow at the outlet 2001-2013

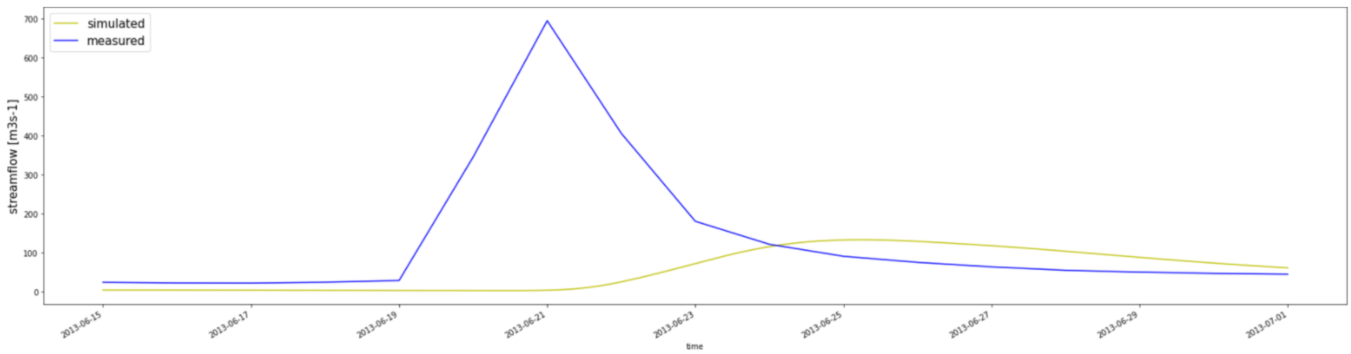

Figure C-10 - Difference between the simulated and measured streamflow for the Elbow during the 2013 flood

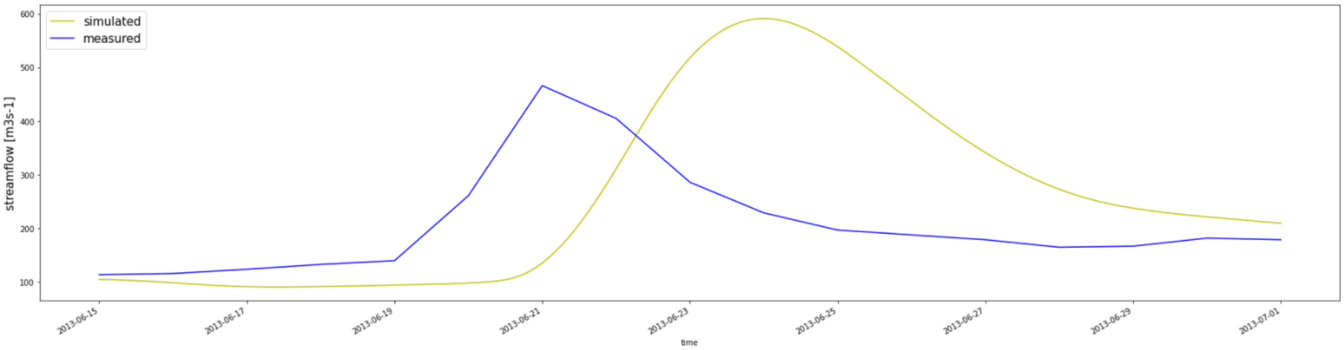

Figure C-11 - Difference between the simulated and measured streamflow for the Bow at Banff during the 2013 flood

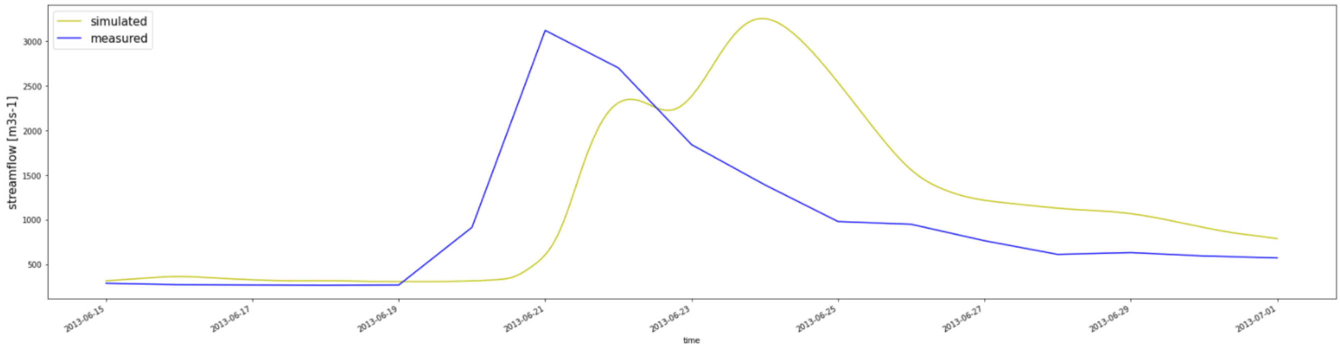

Figure C-12 - Difference between the simulated and measured streamflow at the outlet during the 2013 flood

Input validation

Table C-2 - Precipitation stations

Name of station	Corresponding HRU
BanffCS	71028585
CopUpper	71027942
Elbow Ranger Station	71035477
Lake Louise	71031893
Little Elbow Summit	71036451

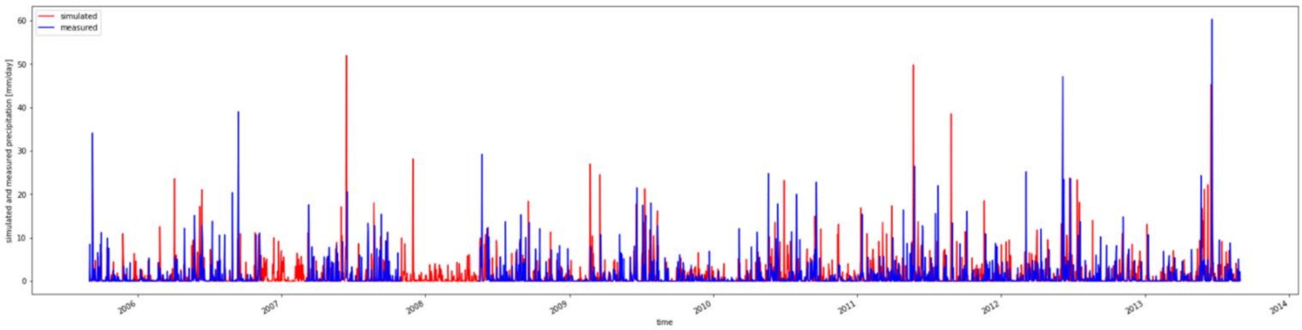

Figure C-13 – WRF simulated and measured precipitation Banff CS

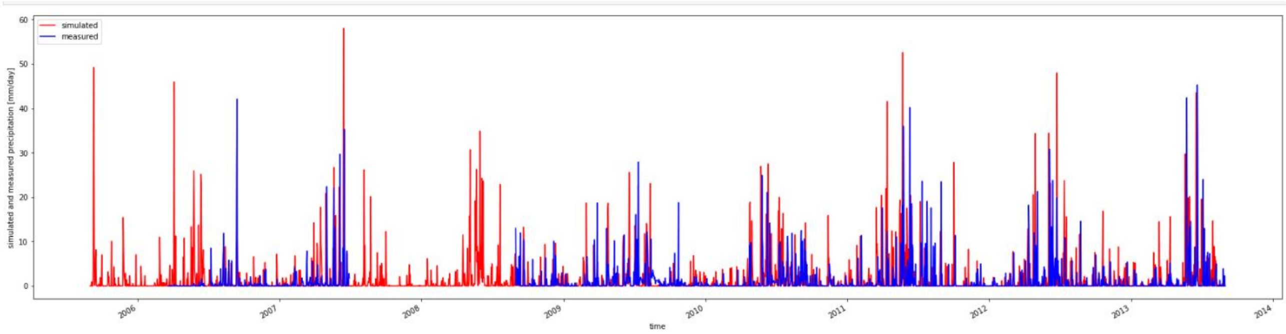

Figure C-14 – WRF simulated and measured precipitation CopUpper

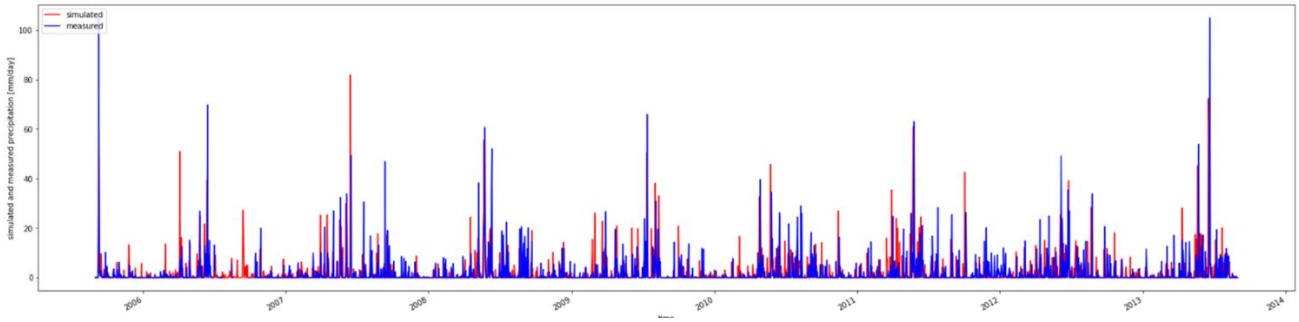

Figure C-15 – WRF simulated and measured precipitation Elbow Ranger Station

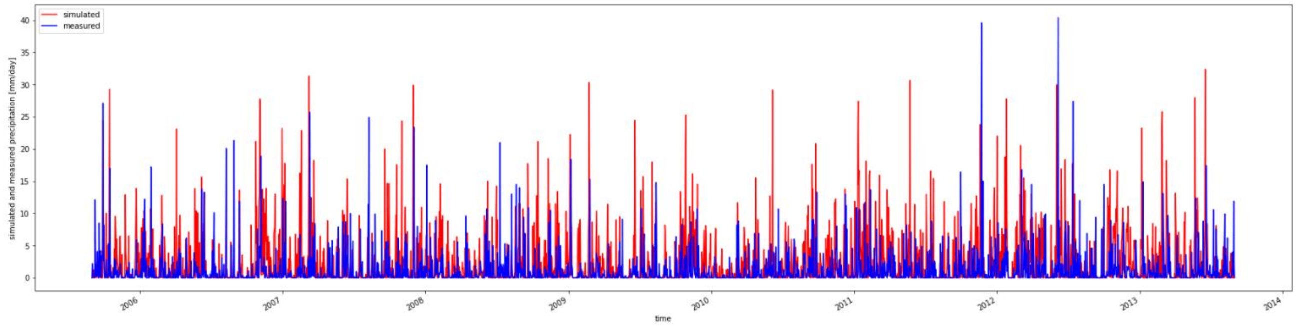

Figure C-16 – WRF simulated and measured precipitation Lake Louise

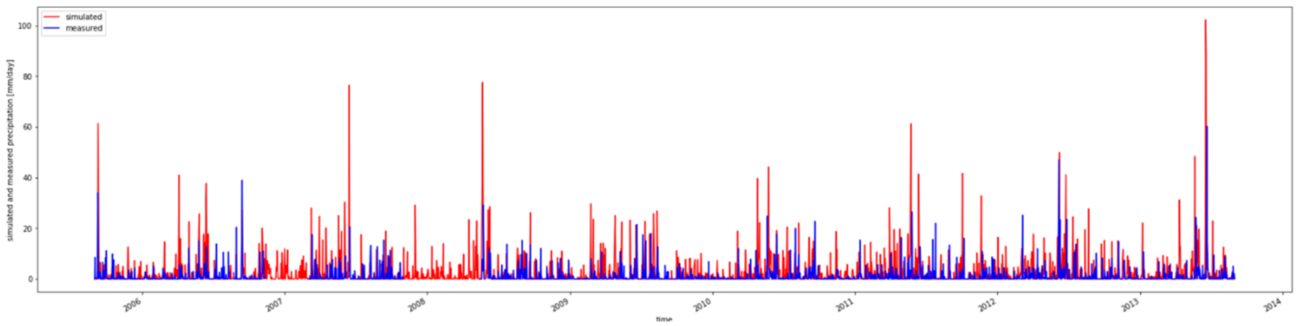

Figure C-17 – WRF simulated and measured precipitation Little Elbow Summit

Appendix D – Selecting representative sub-catchments

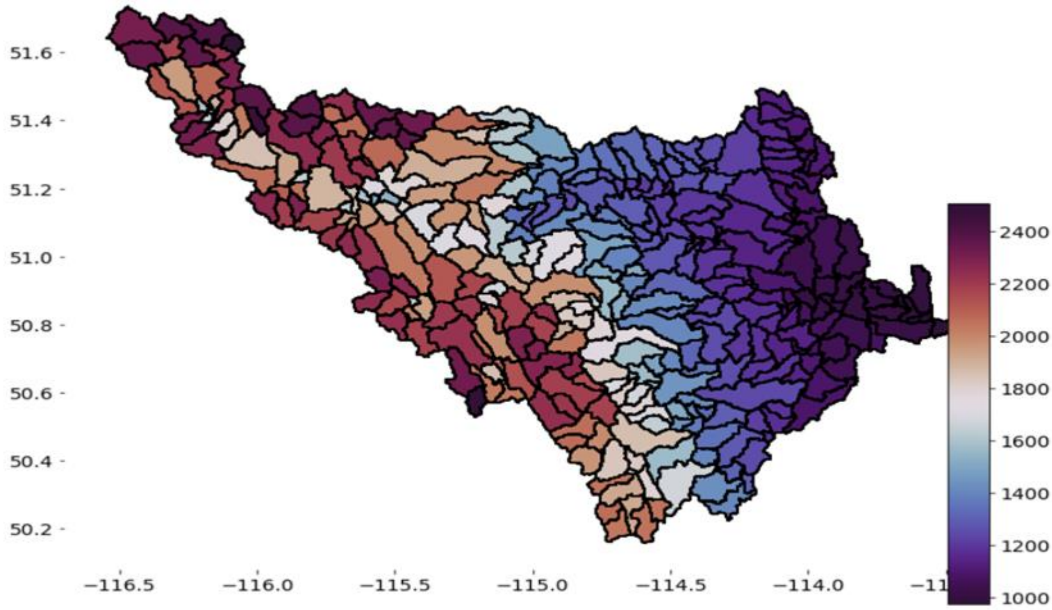

Figure D-1 - Mean elevation sub-catchments [m]

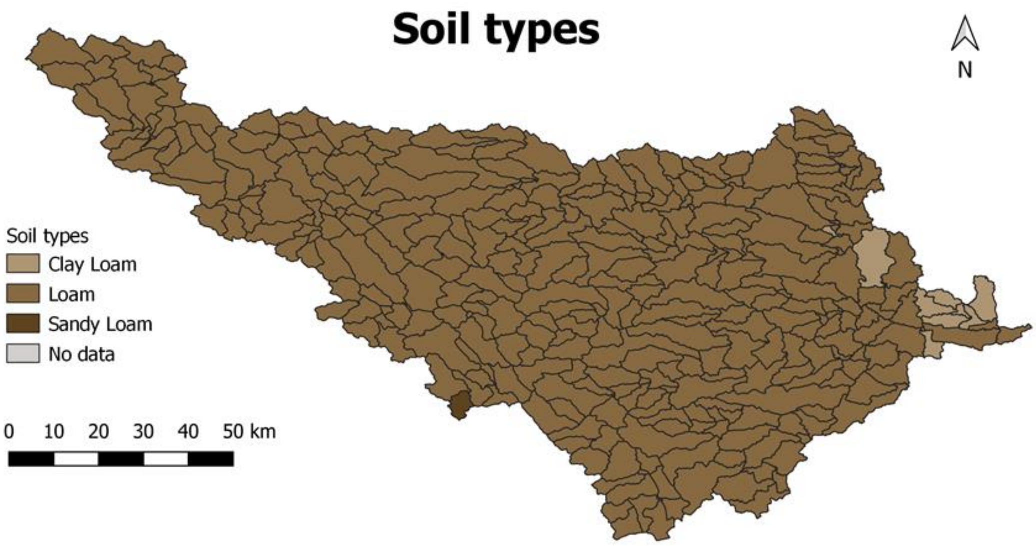

Figure D-2 - Soil types of sub-catchments

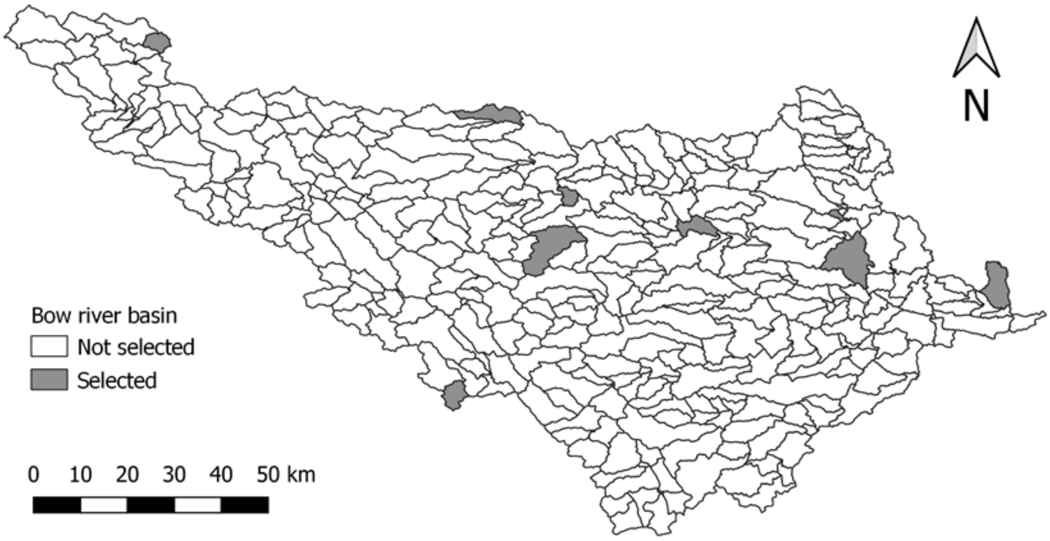

Figure D-3 - Selected sub-catchments

Appendix E – Initial conditions

In Table E-1 an overview of the different variable for which values of initial conditions can be installed is given, in which the snow initial conditions are shaded blue.

Table E-1 - Overview of SUMMA initial conditions (Clark, et al., 2021)

Variable	Data type	Unit	Explanation
nSnow	int	-	The number of snow layers
scalarSnowAlbedo	double	-	Snow albedo for the entire spectral band
scalarSnowDepth	double	m	Total snow depth
scalarSWE	double	kg m^{-2}	Snow water equivalent
dt_init	double	seconds	Length of the initial time sub-step at start of next time interval
scalarCanopyIce	double	kg m^{-2}	Mass of ice on the vegetation canopy
scalarCanopyLiq	double	kg m^{-2}	Mass of liquid water on the vegetation canopy
scalarCanairTemp	double	Pa	Temperature of the canopy air space
scalarCanopyTemp	double	K	Temperature of the vegetation canopy
scalarSfcMeltPond	double	kg m-2	Ponded water caused by melt of the "snow without a layer"
scalarAquiferStorage	double	m	Relative aquifer storage -- above bottom of the soil profile
iLayerHeight	double	m	Height of the layer interface; top of soil = 0
mLayerDepth	double	m	Depth of each layer
layer			
mLayerTemp	double	K	Temperature of each layer
mLayerVolFracIce	double	-	Volumetric fraction of ice in each layer
mLayerVolFracLiq	double	-	Volumetric fraction of liquid water in each layer
mLayerMatricHead	double	m	Matric head of water in the soil

In Table E-2 an overview is given of the start dates of the flood forecasts for the different days of the flood curve and all lead times. Dates that differ from the exact date for that lead time, because otherwise it would not be able to start the hindcast, are shaded orange. The hindcasts could not be started at that date because of a bug in SUMMA. One is not able to start a simulation on a day where there is canopy ice and the temperature is also above 0, therefore, some hindcasts are started a day earlier or later.

Table E-2 - Start dates of the flood hindcasts for the different days of the flood curve and all lead times

	22 June	**23 June**	**24 June**	**25 June**	**26 June**
Lead time					
1 day	21 June	22 June	23 June	24 June	25 June
2 days	20 June	21 June	22 June	23 June	24 June
3 days	19 June	20 June	21 June	22 June	23 June
4 days	18 June	19 June	20 June	21 June	22 June
5 days	17 June	18 June	19 June	20 June	21 June
6 days	16 June	17 June	18 June	19 June	20 June
1 week	15 June	16 June	17 June	18 June	19 June
2 weeks	8 June	9 June	10 June	11 June	12 June
3 weeks	1 June	2 June	3 June	3 June	5 June
4 weeks	25 May	25 May	27 May	28 May	28 May
8 weeks	27 April	28 April	29 April	30 April	1 May

Appendix F – Thresholds hindcast assessment

Table F-1 - Thresholds used for calculation dichotomous skill scores (The city of Calgary, 2021)

Description of flood impact	Annual chance of occurrence	Bow at Calgary	Elbow at Calgary
Low normal seasonal flows	99.9 %	70 m^3s^{-1}	30 m^3s^{-1}
Pathways may be impacted	>50 %	250 m^3s^{-1}	40 m^3s^{-1}
Widespread basement flooding and evacuation begins	13-17 %	700 m^3s^{-1}	170 m^3s^{-1}
Widespread evacuation	3-5 %	1500 m^3s^{-1}	275 m^3s^{-1}
Hundred year flood rate	1 %	2020 m^3s^{-1}	803 m^3s^{-1}

Table F-2 - Thresholds Reliability diagram

	Carseland dam	Bow at Calgary	Elbow at Calgary	Lake Louise	Sheep river
Percentiles used	[0.8, 1]	[0.75, 1]	[0.75, 1]	[0.9, 1]	[0.6, 1]
Low threshold [$m^3\ s^{-1}$]	565	361	27	122	57
High threshold [$m^3\ s^{-1}$]	3253	2593	133	324	1752

Table F-3 - Thresholds qualitative assessment forecasts

	Outlet	Bow Calgary	Elbow Calgary	Bow upstream	Sheep river
50%	249	185	13	45	34
98%	1242	893	101	168	313
Maximum before 2013	1898	1634	118	201	797
Maximum 2013	3253	2593	133	324	1752

Appendix G – Hydrological differences between the sub-catchments

Snow water equivalent

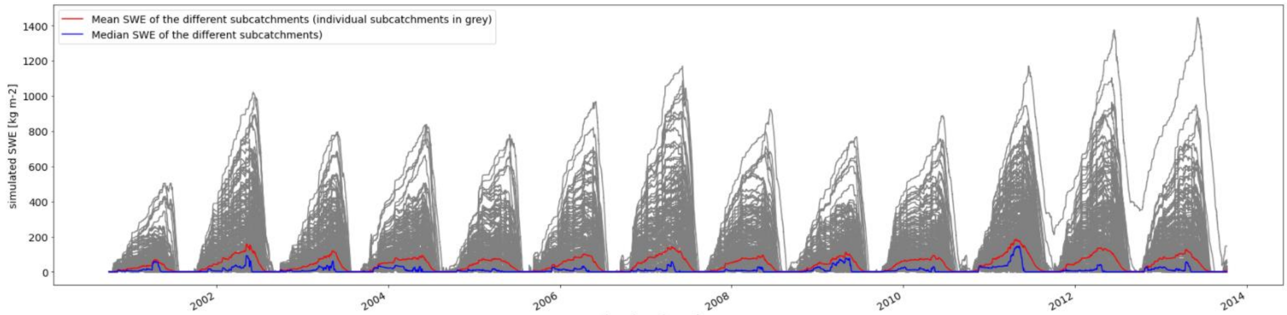

Figure G-1 - Snow water equivalent of all sub-catchments

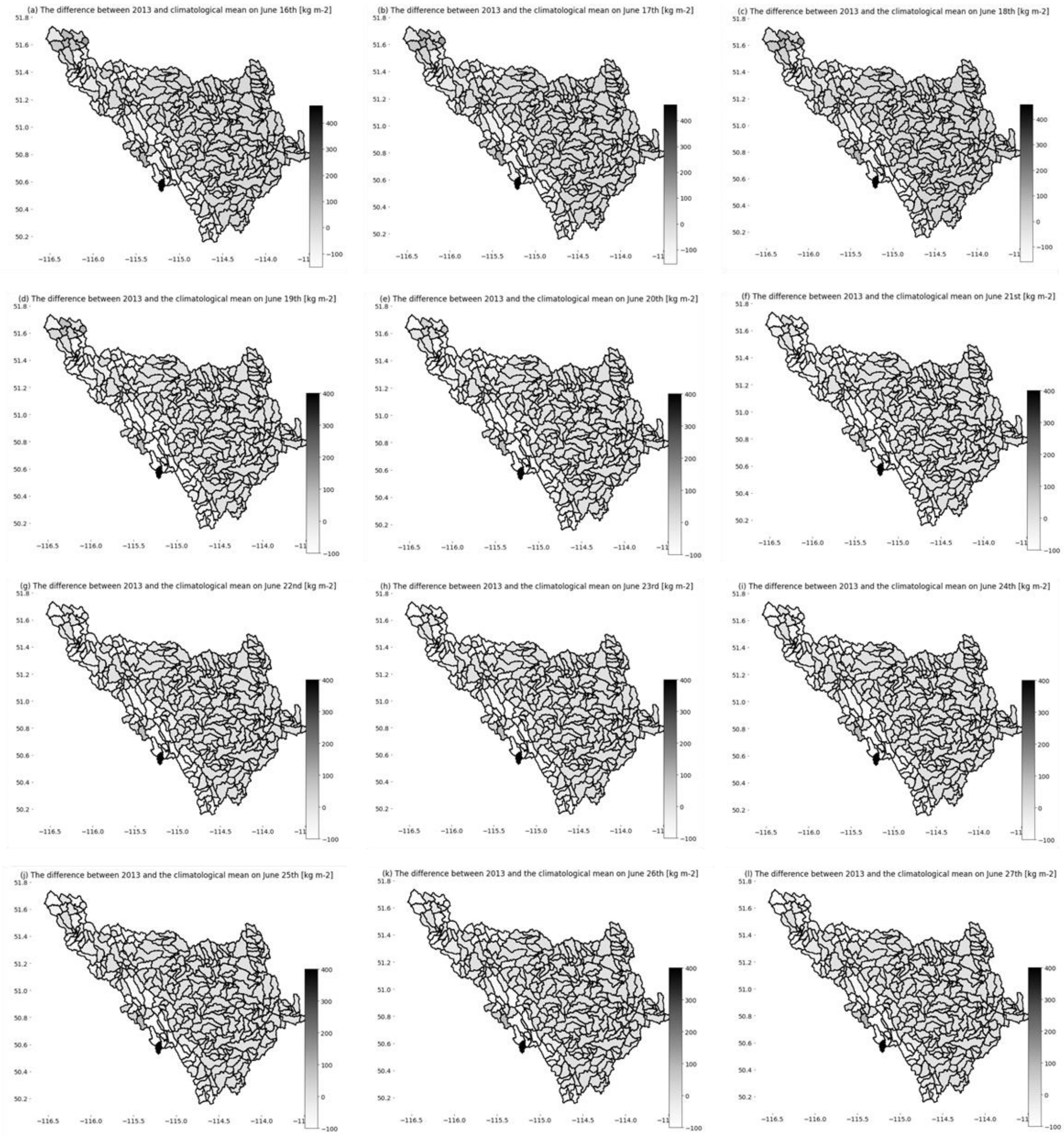

Figure G-2 - Difference in SWE between 2013 and climatological mean for the different sub-catchments

Temperature

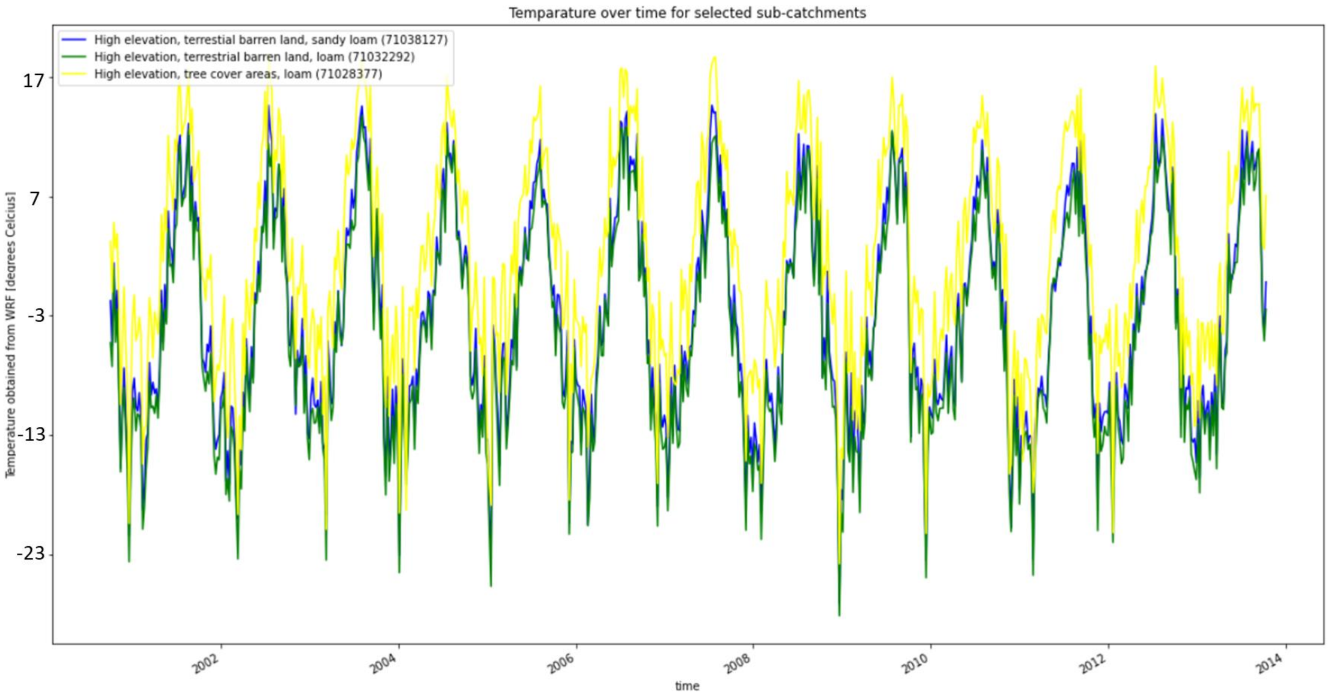

Figure G-3 - Temperature over time for the selected sub-catchments with high elevation

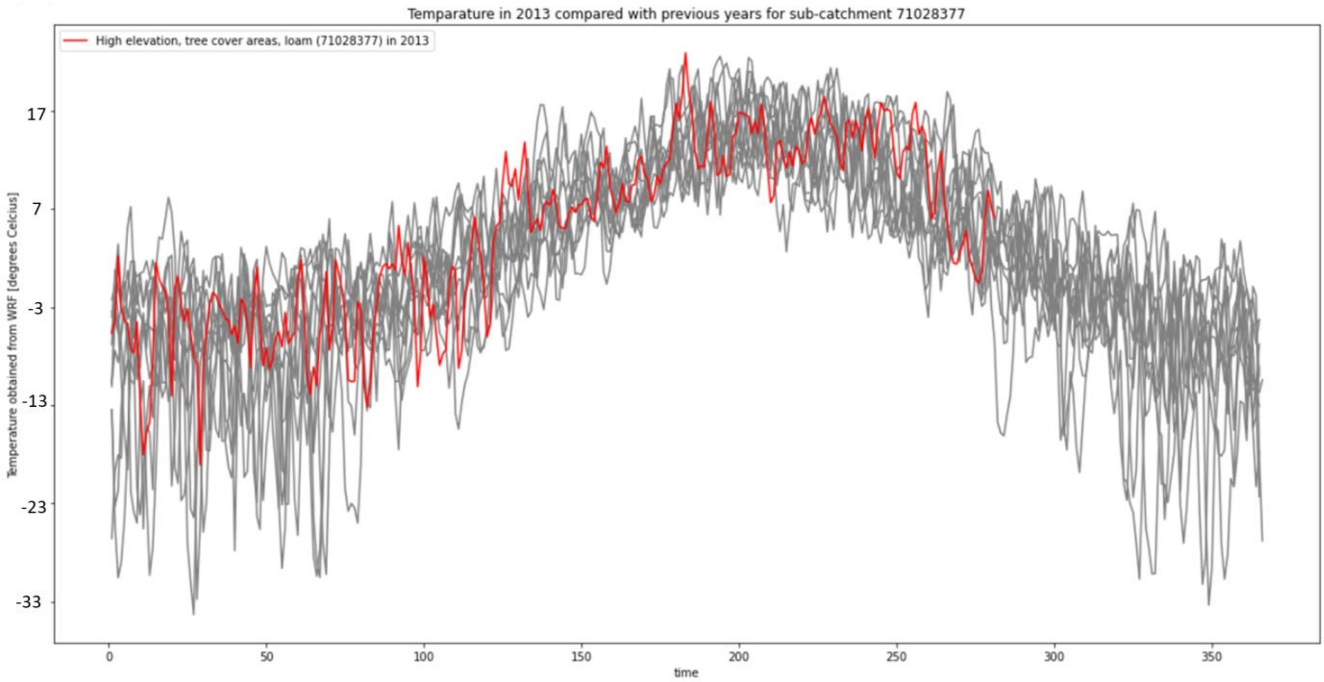

Figure G-4 - Temperature in 2013 compared with previous years for sub-catchment 71028377

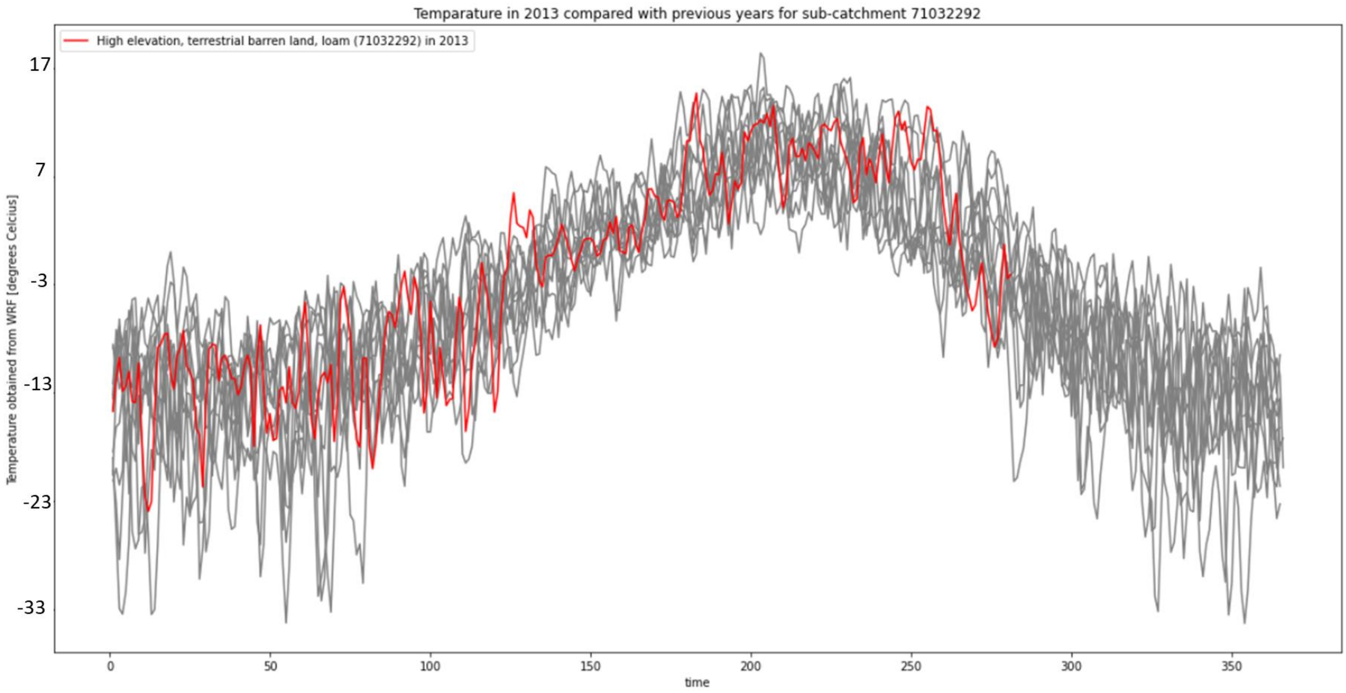

Figure G-5 - Temperature in 2013 compared with previous years for sub-catchment 71032292

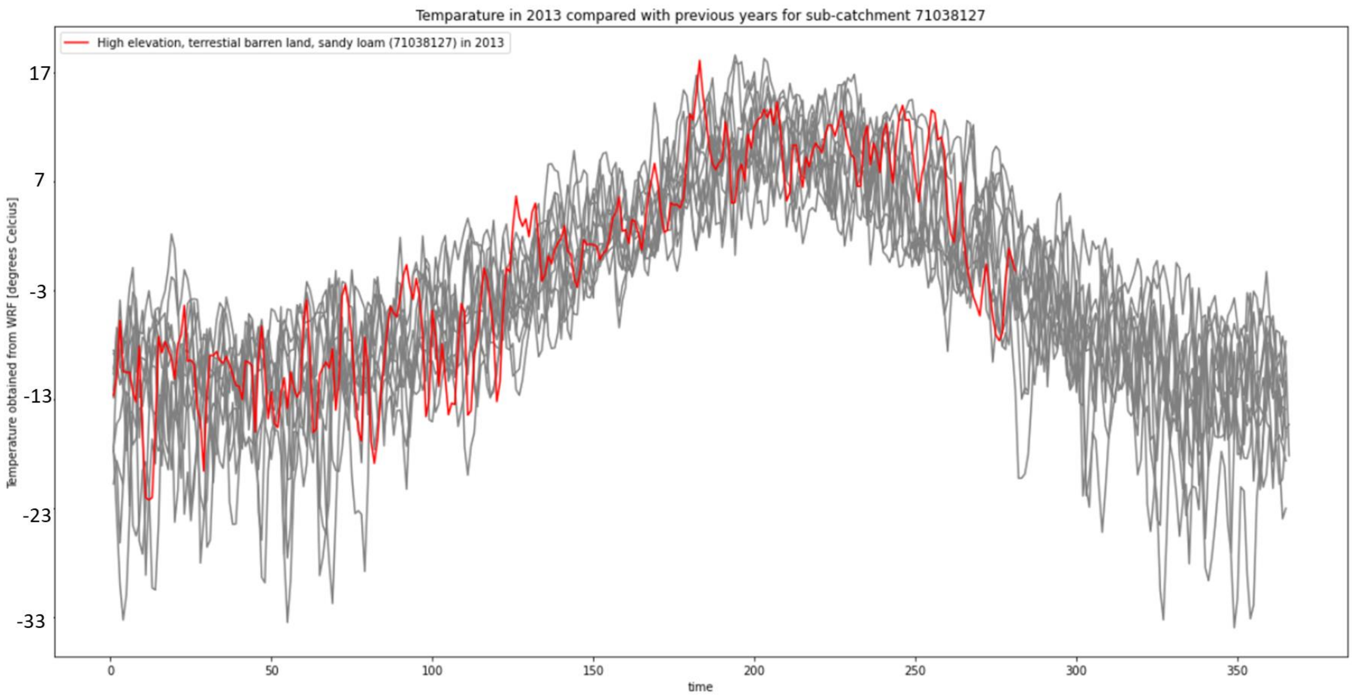

Figure G-6 - Temperature in 2013 compared with previous years for sub-catchment 71038127

Streamflow

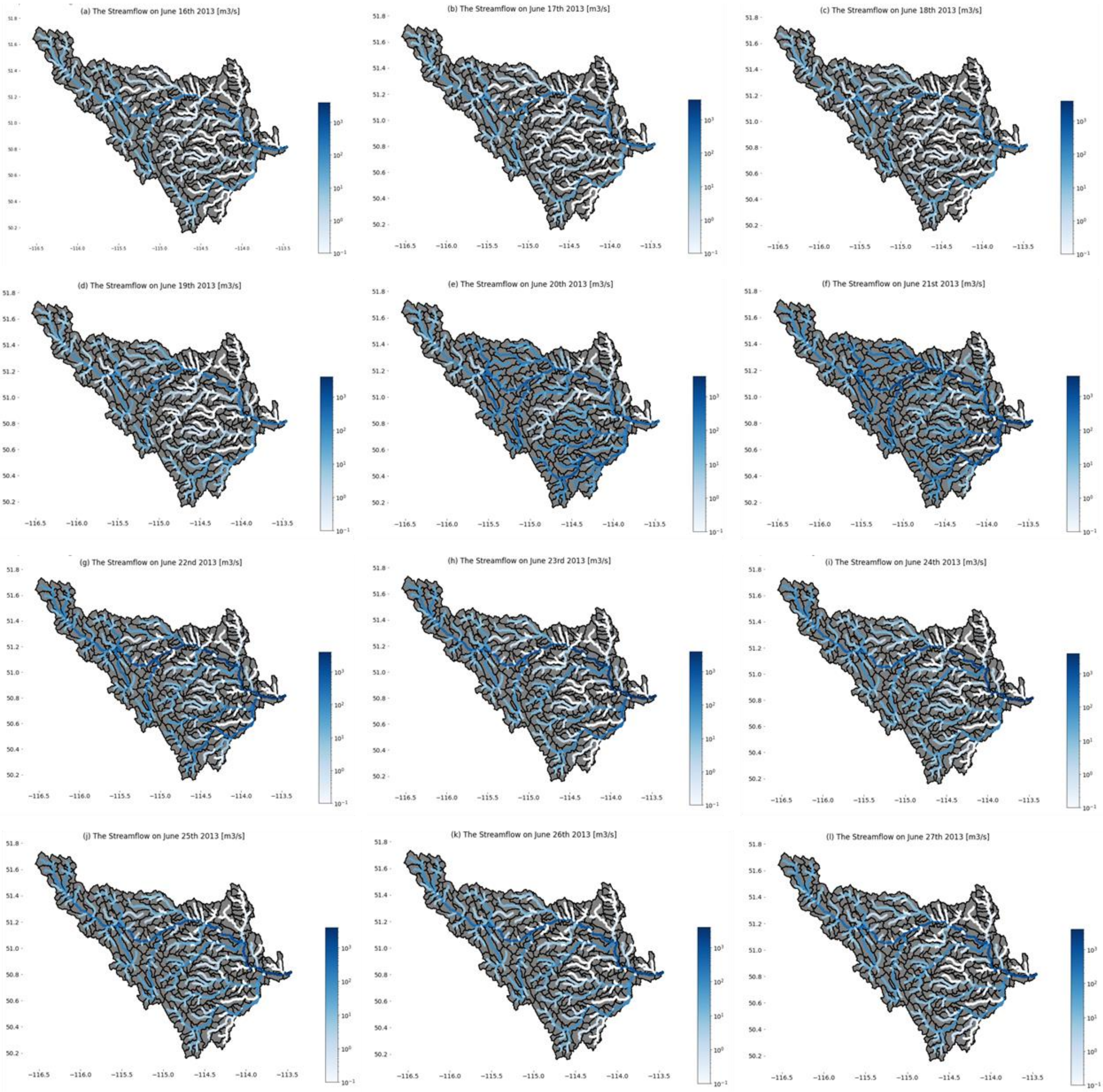

Figure G-7 - Streamflow in the different sub-catchments of the study area

Appendix H – Visual representation of flood forecasts

In Figure H-1 up and until H-25 a visual representation of the flood forecasts is given. A different plot is made for each day of the flood curve and for each of the selected locations. In the plot the range between the maximum forecasted value for that specific day of the flood curve at that location of all the ensemble members and the minimum forecasted value for that specific day of the flood curve at that location of all the ensemble members is presented with a blue line for each lead time. The mean value of the ensemble members is indicated by a blue dot. The 'perfect simulation' of the discharge, as obtained from model spin-up in the first part of this research, is indicated with a green line. The maximum discharge that occurs in 2013 for that location is indicated with a black line. The grey box represents the range between the 50th percentile and the maximum discharge of the period 2001-2012 for the period of 15 May until 15 July. Which means that if the forecasts in the figure are in the grey box or above, the forecasts are above average for the general flood period. There is opted for this period since this is the general period in which seasonal floods occur (The city of Calgary, 2021).

22nd of June

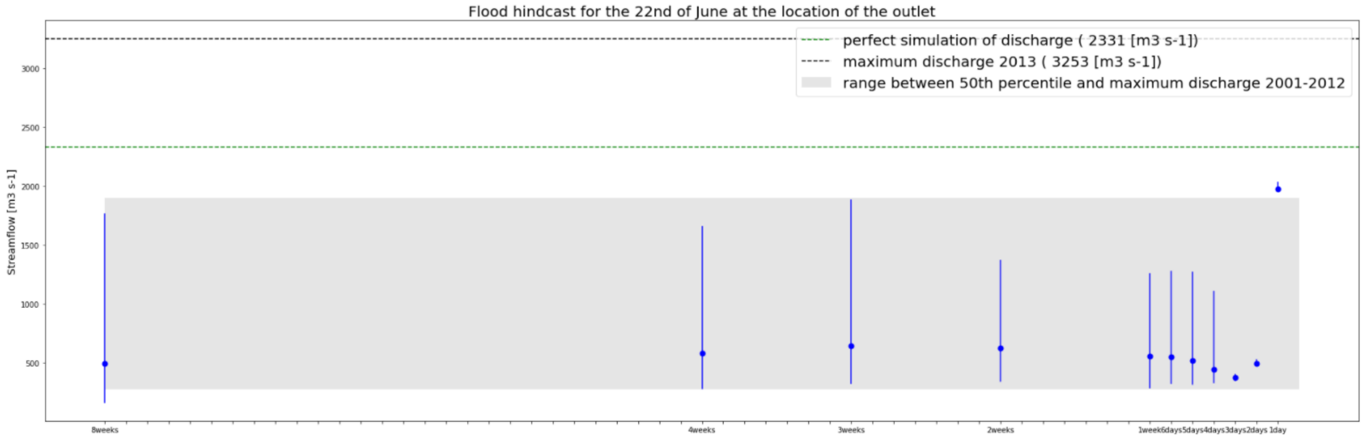

Figure H-1 - Flood hindcasts for the 22nd of June at the Carseland dam for all lead times

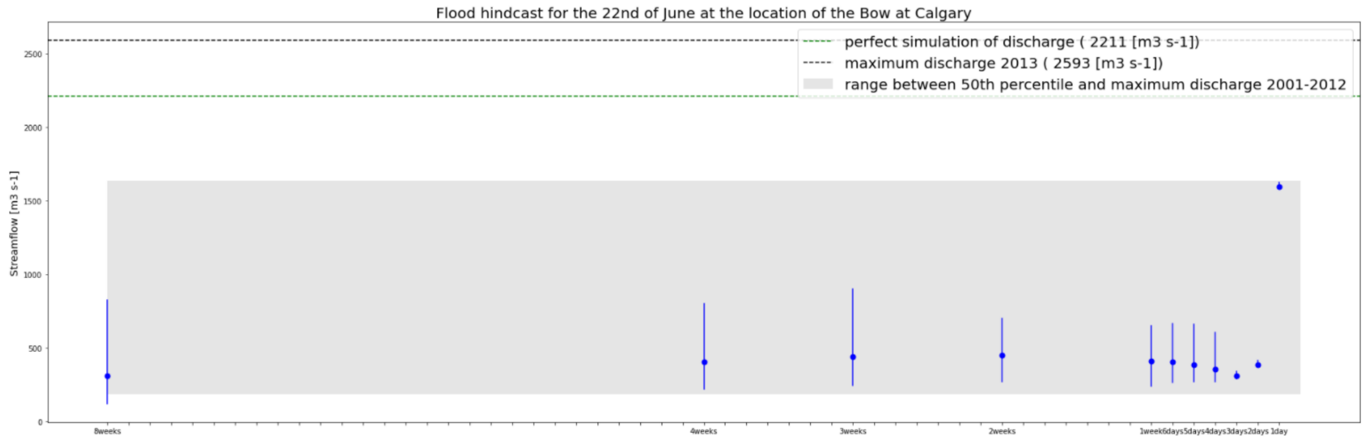

Figure H-2 - Flood hindcasts for the 22nd of June at the location of the Bow at Calgary for all lead times

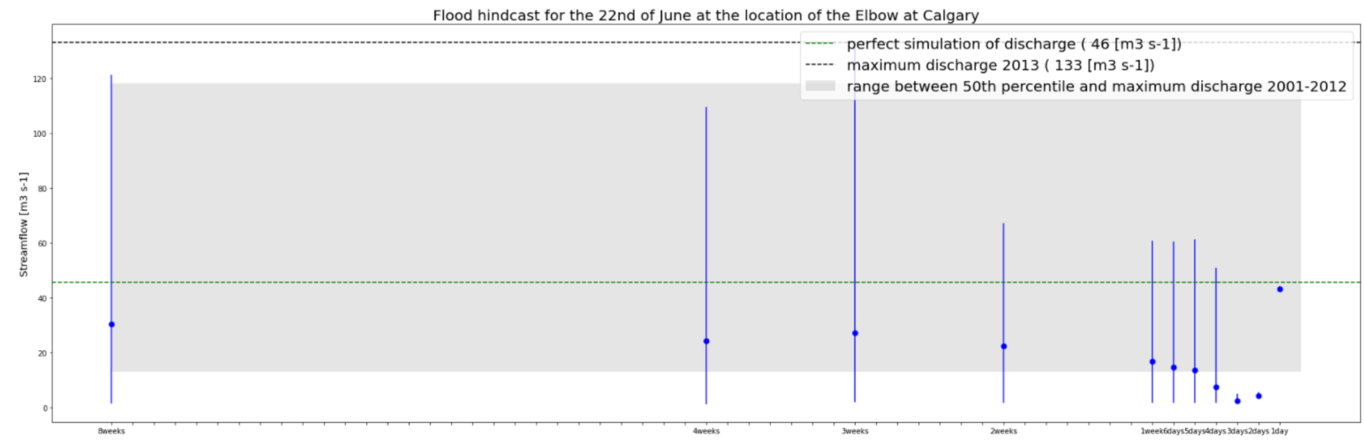

Figure H-3 - Flood hindcasts for the 22nd of June at the location of the Elbow at Calgary for all lead times

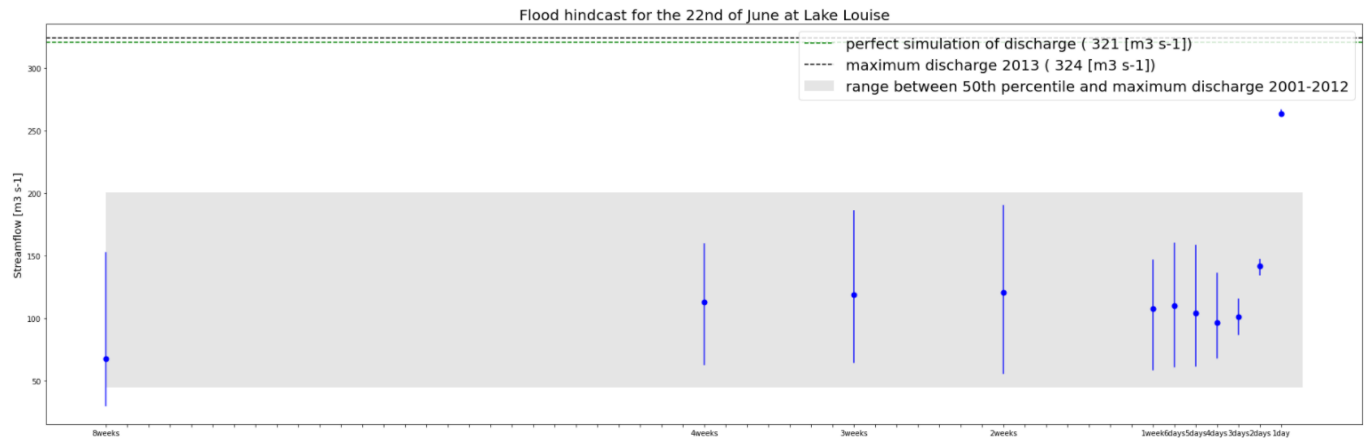

Figure H-4 - Flood hindcasts for the 22nd of June at Lake Louise for all lead times

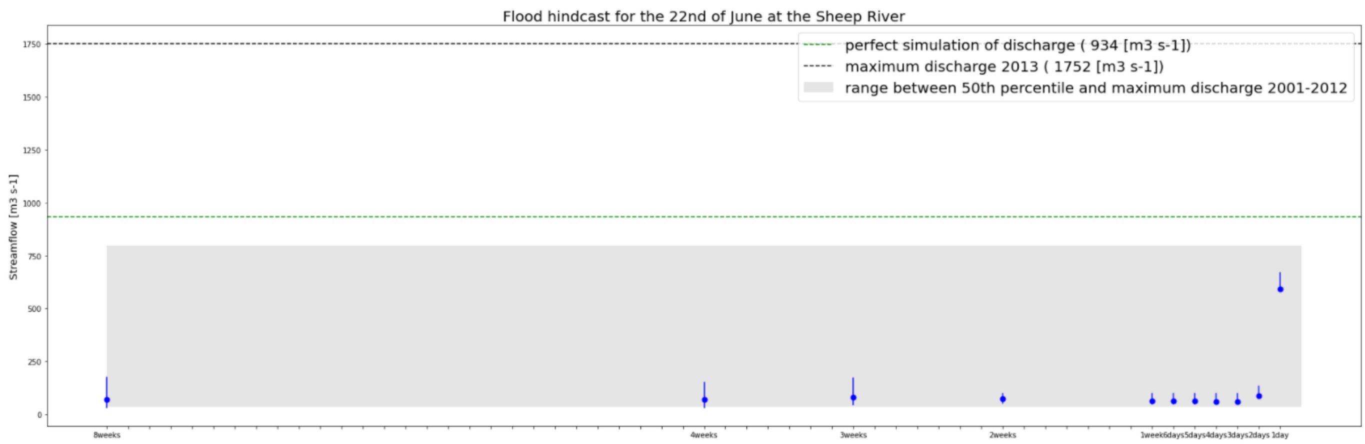

Figure H-5 - Flood hindcasts for the 22nd of June at the Sheep River for all lead times

23rd of June

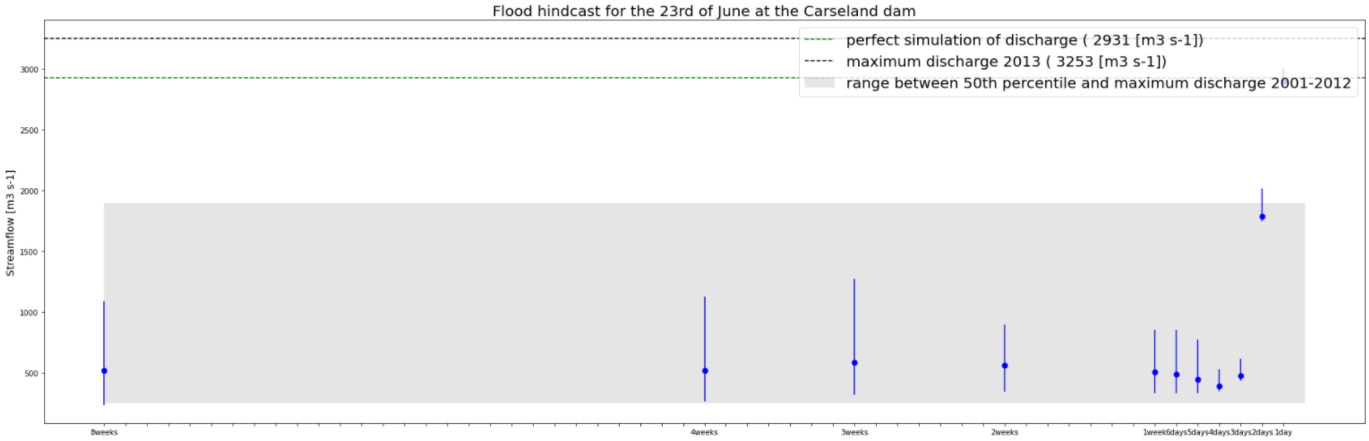

Figure H-6 - Flood hindcasts for the 23rd of June at the Carseland dam for all lead times

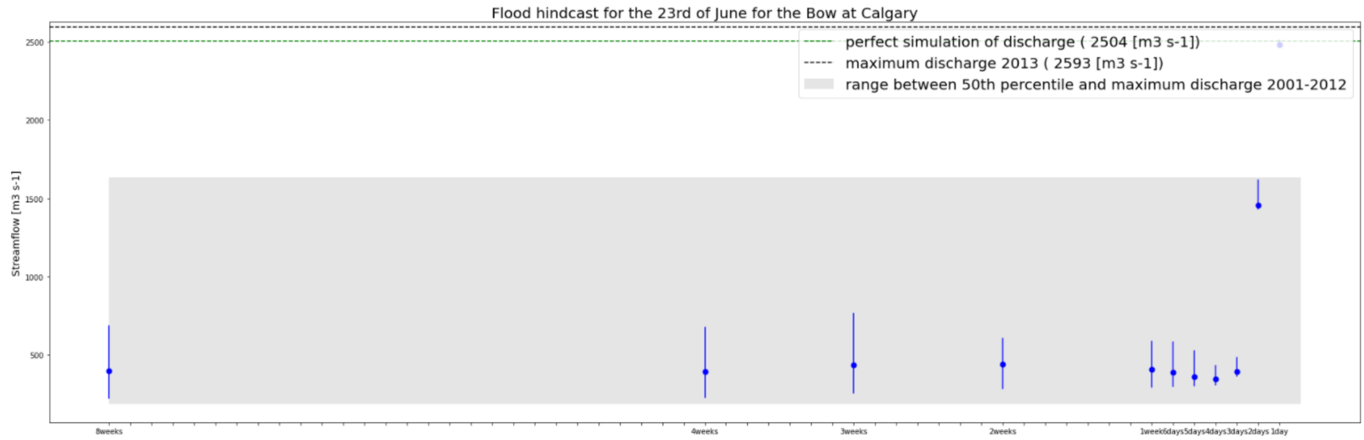

Figure H-7 - Flood hindcasts for the 23rd of June for the Bow at Calgary for all lead times

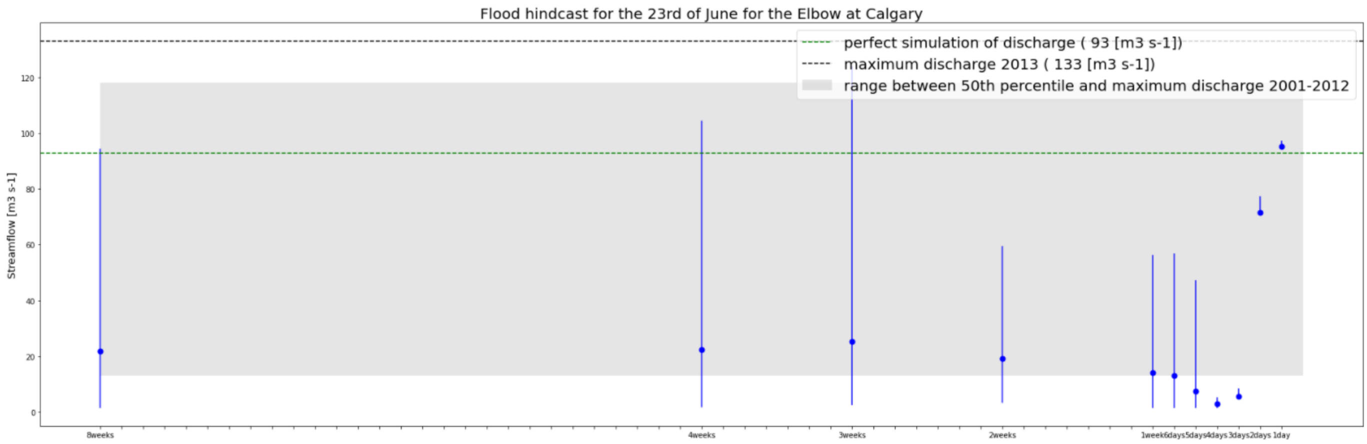

Figure H-8 - Flood hindcasts for the 23rd of June for the Elbow at Calgary for all lead times

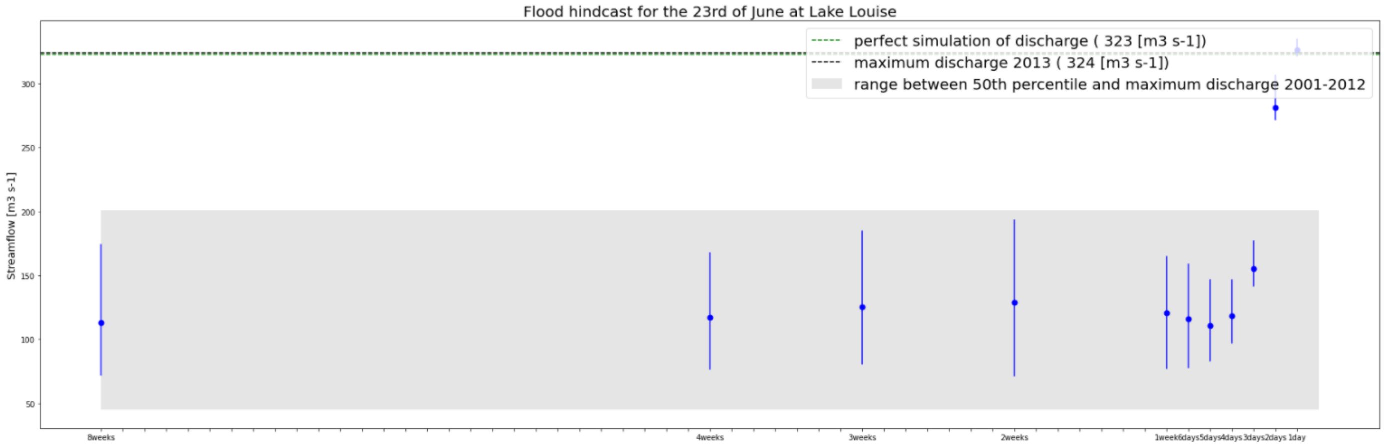

Figure H-9 - Flood hindcasts for the 23rd of June at Lake Louise for all lead times

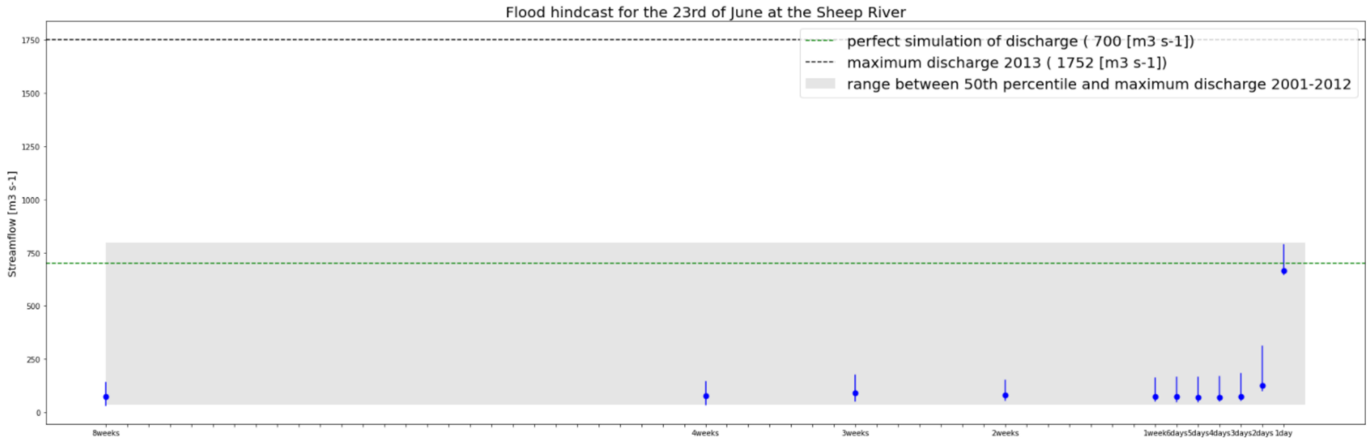

Figure H-10 - Flood hindcasts for the 23rd of June at the Sheep River for all lead times

24th of June

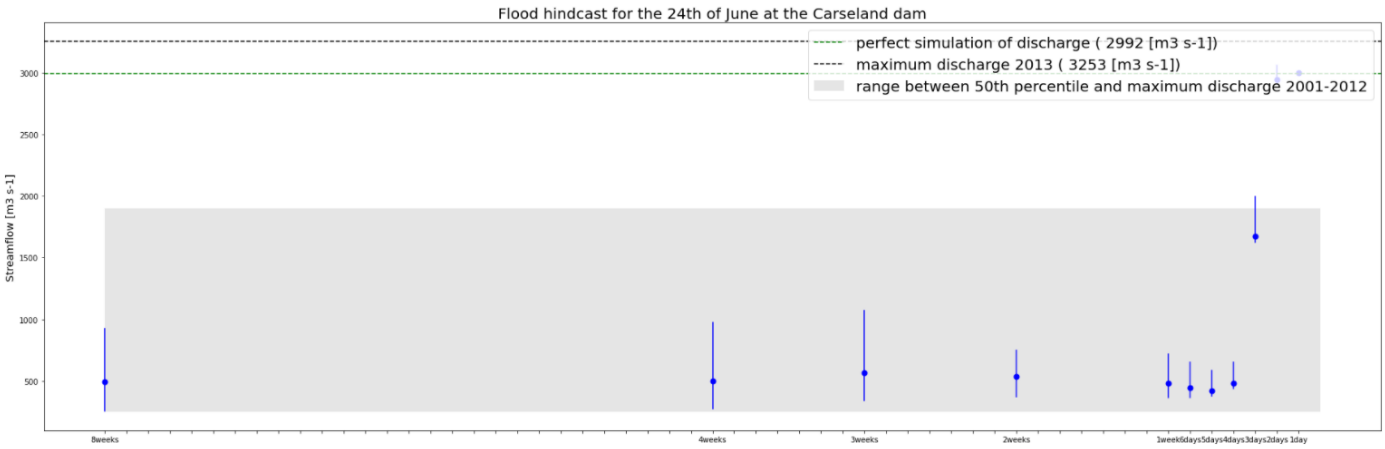

Figure H-11 - Flood hindcasts for the 24th of June at the Carseland dam for all lead times

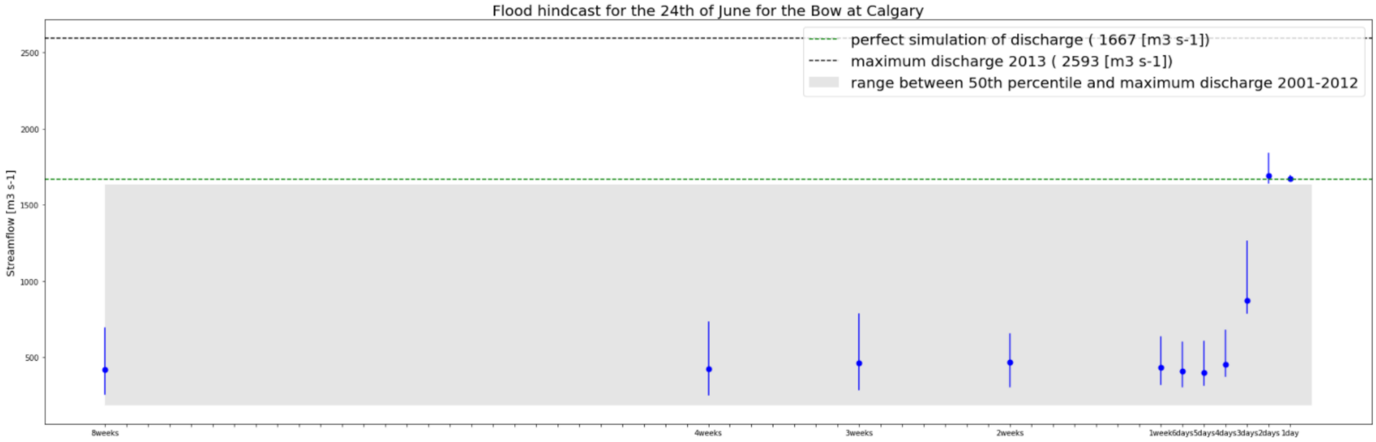

Figure H-12 - Flood hindcasts for the 24th of June for the Bow at Calgary for all lead times

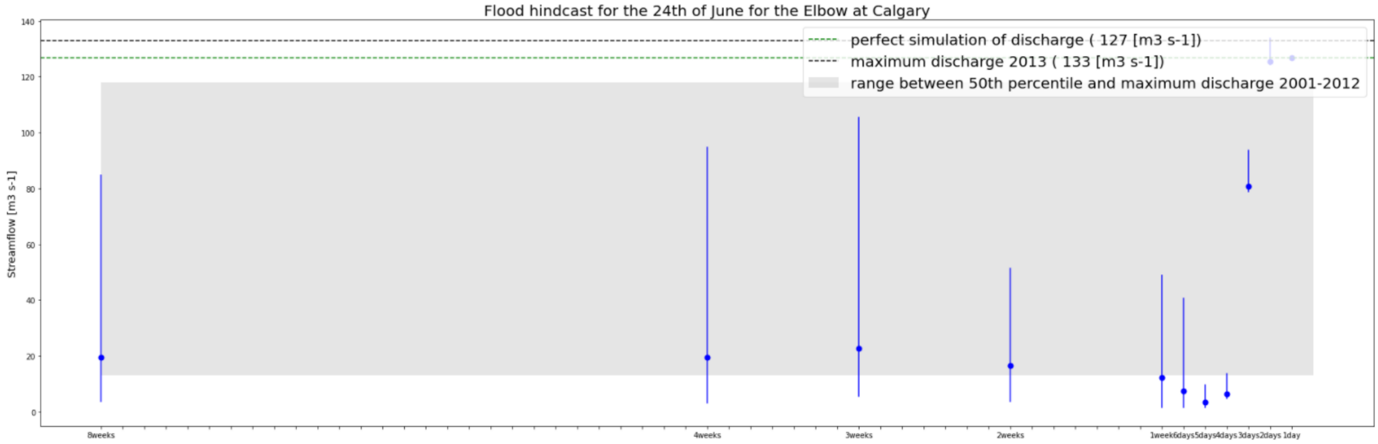

Figure H-13 - Flood hindcasts for the 24th of June for the Elbow at Calgary for all lead times

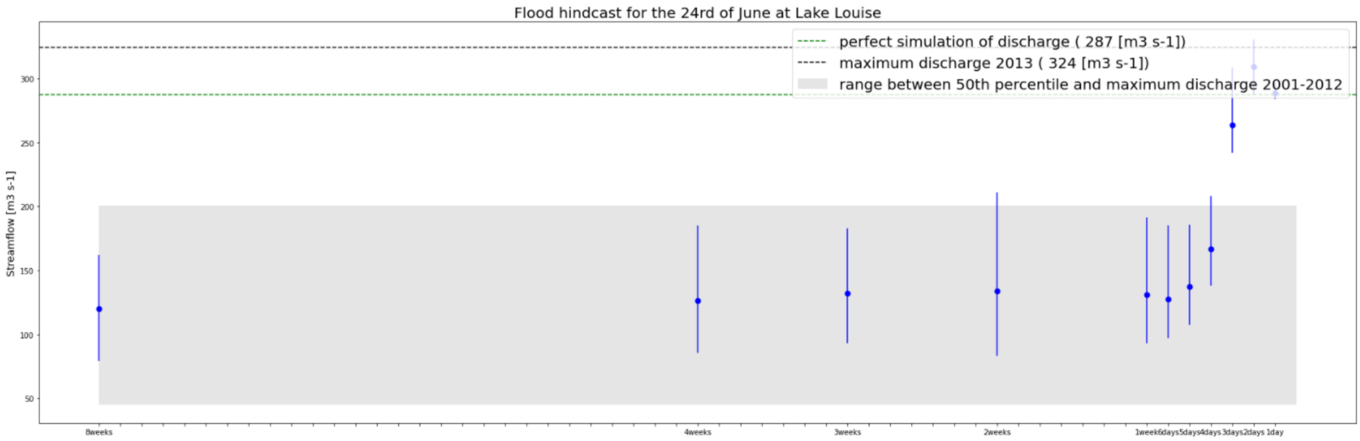

Figure H-14 - Flood hindcasts for the 24th of June at Lake Louise for all lead times

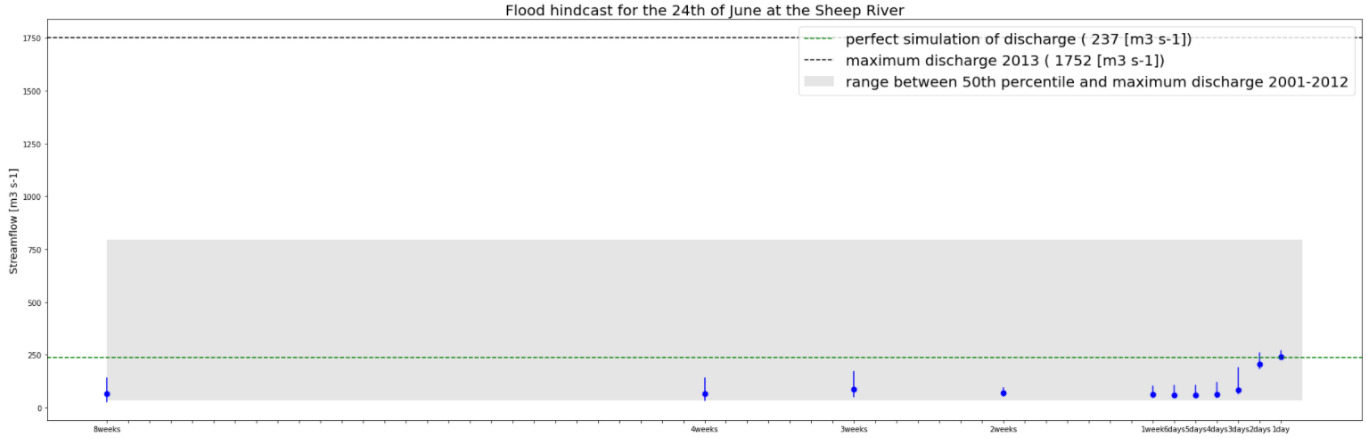

Figure H-15 - Flood hindcasts for the 24th of June at the Sheep River for all lead times

25th of June

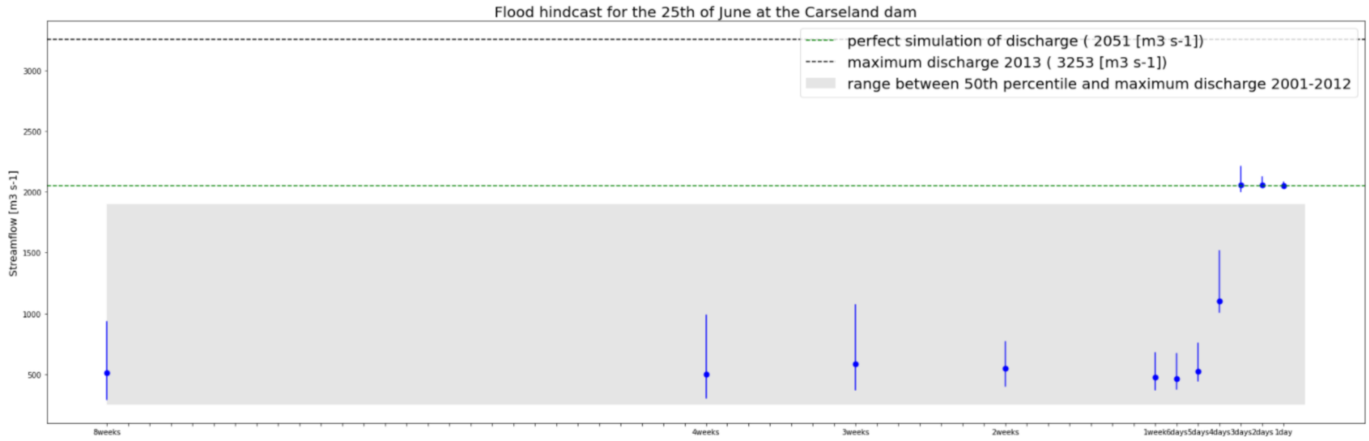

Figure H-16 - Flood hindcast for the 25th of June at the Carseland dam for all lead times

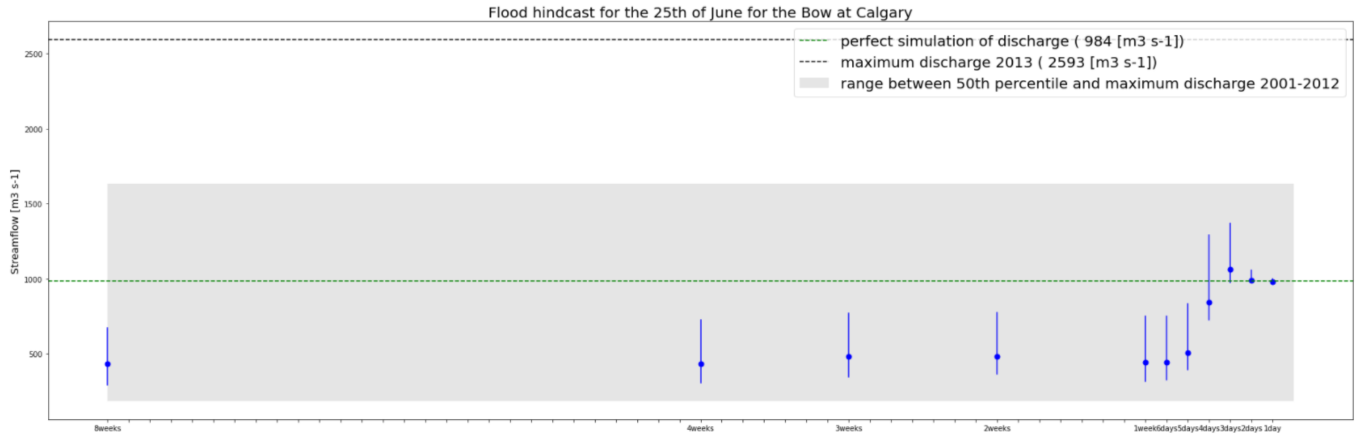

Figure H-17 - Flood hindcasts for the 25th of June for the Bow at Calgary for all lead times

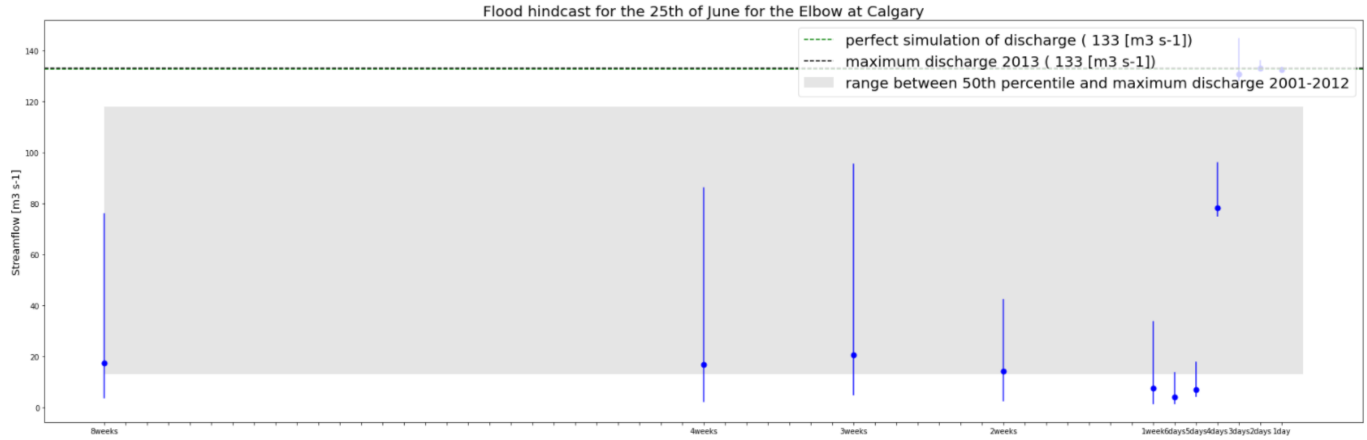

Figure H-18 - Flood hindcasts for the 25th of June for the Elbow at Calgary for all lead times

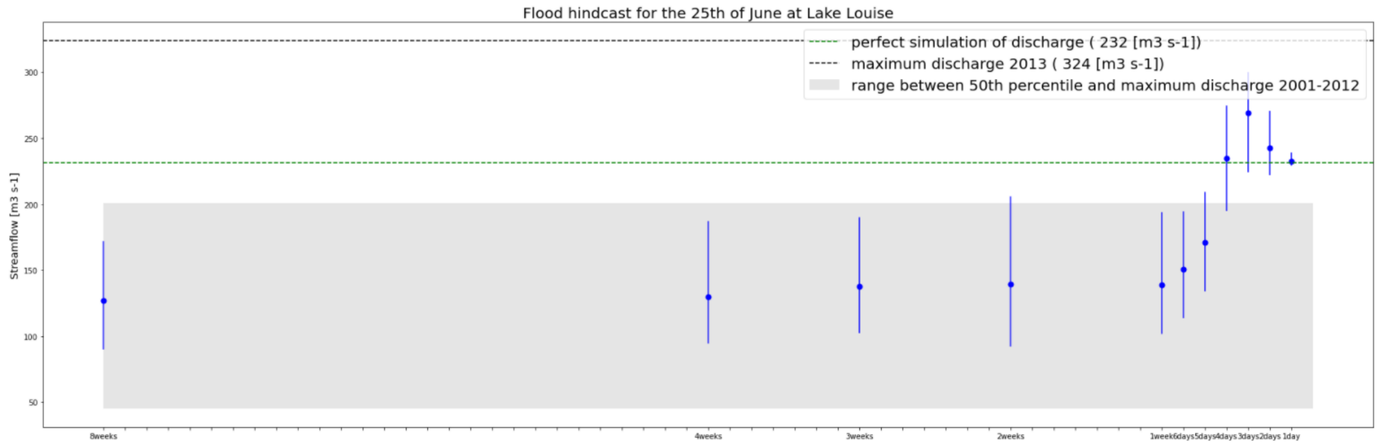

Figure H-19 - Flood hindcasts for the 25th of June at Lake Louise for all lead times

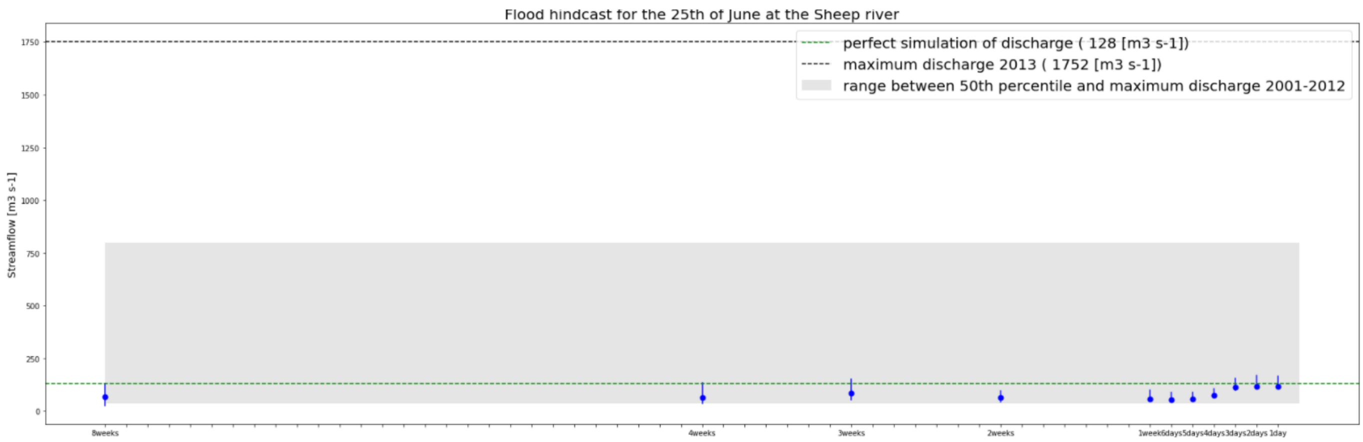

Figure H-20 - Flood hindcasts for the 25th of June at the Sheep River for all lead times

26th of June

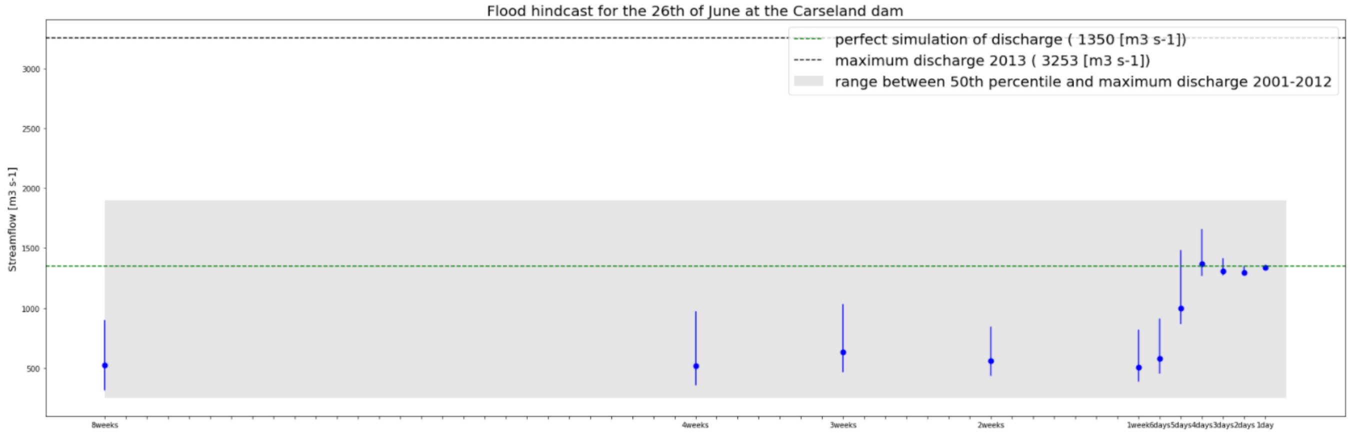

Figure H-21 - Flood hindcast for the 26th of June at the Carseland dam for all lead times

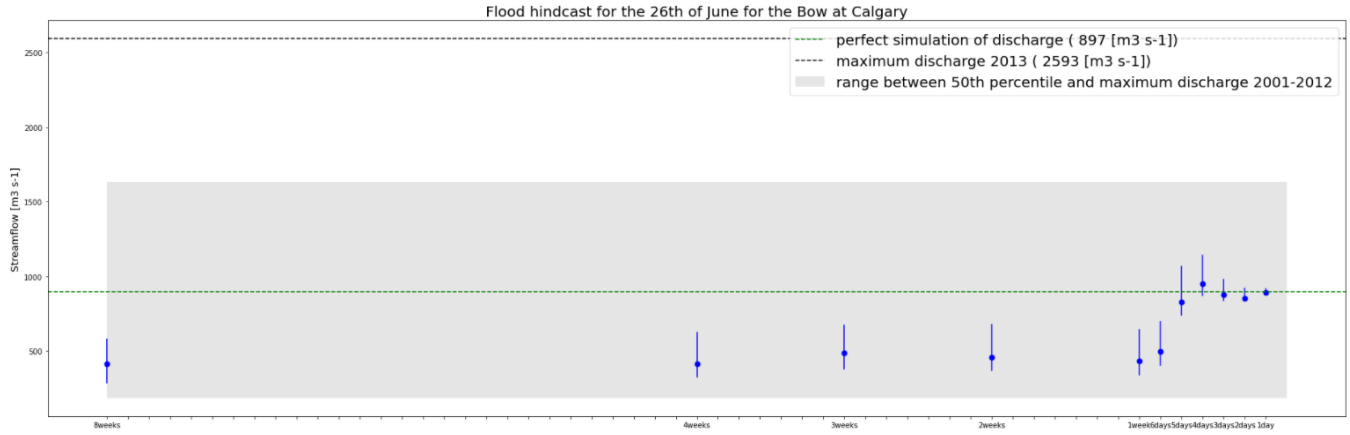

Figure H-22 - Flood hindcasts for the 26th of June for the Bow at Calgary for all lead times

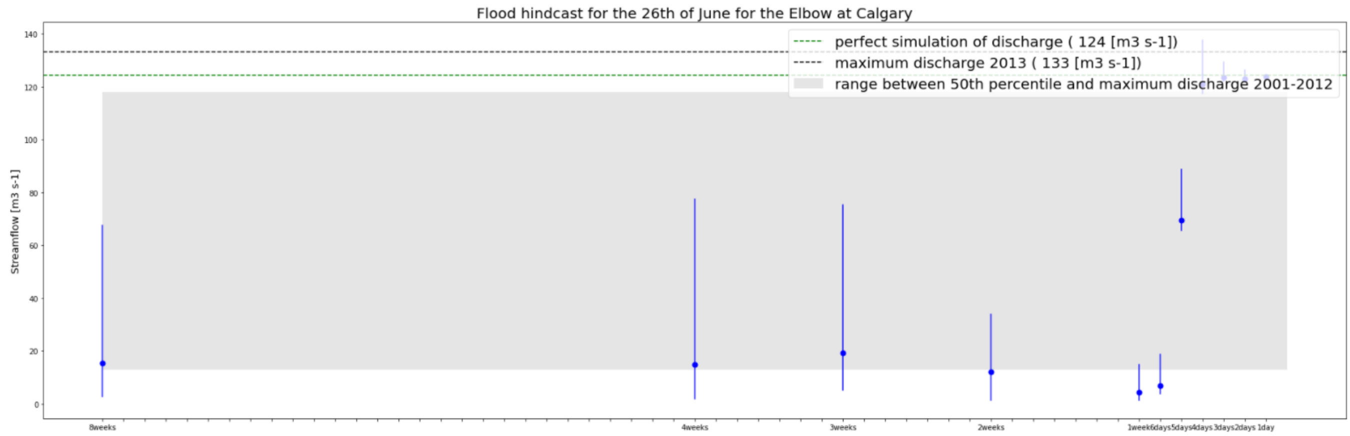

Figure H-23 - Flood hindcasts for the 26th of June for the Elbow at Calgary for all lead times

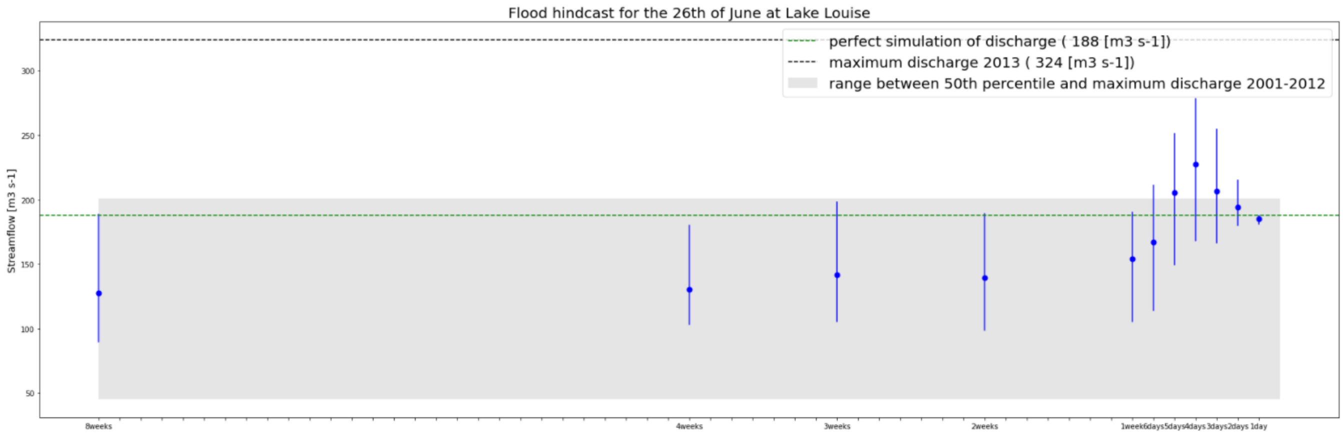

Figure H-24 - Flood hindcasts for the 26th of June at Lake Louise for all lead times

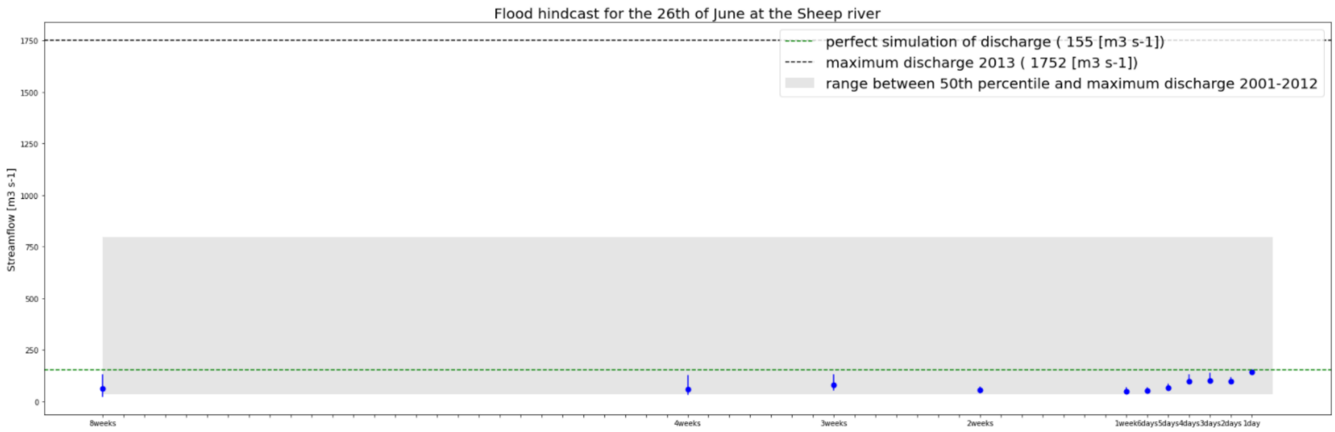

Figure H-25 - Flood hindcasts for the 26th of June at the Sheep River for all lead times

Appendix I – Continuous Ranked Probability (Skill) Score (CRP(S)S)

CRPS

Table I-1 - CRPS at the Carseland dam [m^3s^{-1}]

	Date	22 June	23 June	24 June	25 June	26 June
Lead time						
1 day		314.18	442.66	476.68	171.48	79.92
2 days		1794.83	638.57	402.56	169.62	31.08
3 days		1910.2	1953.49	821.96	139.73	33.78
4 days		1770.16	2041.23	2022.21	738.66	53.88
5 days		1658.94	1943.36	2088.45	1330.1	241.11
6 days		1623.98	1880.82	2040.84	1391.13	638.52
1 week		1609.33	1856.17	1984.84	1360.34	714.21
2 weeks		1518.6	1786.37	1921.85	1273.69	635.48
3 weeks		1461.09	1728.02	1858.54	1210.73	544.95
4 weeks		1546.32	1810.36	1931.75	1294.52	666.21
8 weeks		1623.48	1820.17	1939.47	1284.63	655.18

Table I-2 – CRPS at Calgary for the Bow river [m^3s^{-1}]

	Date	22 June	23 June	24 June	25 June	26 June
Lead time						
1 day		577.87	366.56	99.89	26.83	62.55
2 days		1787.92	642.15	85.68	29.99	17.14
3 days		1866.65	1702.98	662.55	63.97	35.63
4 days		1783.08	1756.67	1089.81	113.21	94.66
5 days		1737.5	1723.67	1145.67	398.52	13.32
6 days		1713.61	1679.7	1128.99	468.74	290.48
1 week		1702.19	1658.79	1090.71	455.99	353.99
2 weeks		1653.07	1616.0	1055.15	415.56	324.01
3 weeks		1640.87	1604.06	1046.29	407.94	292.46
4 weeks		1688.88	1652.26	1090.15	462.02	369.72
8 weeks		1776.07	1654.09	1100.9	464.87	368.64

Table I-3 - CRPS at Calgary for the Elbow river [m^3s^{-1}]

	Date	22 June	23 June	24 June	25 June	26 June
Lead time						
1 day		3.25	13.03	17.47	18.11	17.06
2 days		42.07	10.26	15.16	18.17	15.98
3 days		43.29	76.00	27.89	14.18	15.82
4 days		35.39	78.24	102.24	35.27	11.54
5 days		28.98	69.64	104.85	106.61	36.00
6 days		27.62	60.64	97.60	109.19	98.71
1 week		25.25	59.56	90.09	103.31	101.06
2 weeks		18.45	53.38	85.41	94.200	89.82
3 weeks		19.89	48.72	74.79	83.50	78.75
4 weeks		20.58	50.61	79.07	88.36	83.70
8 weeks		15.13	50.63	79.71	88.29	83.74

Table I-4 – CRPS at Lake Louise [m^3s^{-1}]

	Date	22 June	23 June	24 June	25 June	26 June
Lead time						
1 day		17.31	24.95	20.83	11.53	1.97
2 days		138.2	15.28	38.11	19.99	8.38
3 days		177.58	139.82	5.47	44.94	19.28
4 days		178.66	176.21	91.12	9.91	40.86
5 days		169.99	181.18	120.55	37.05	17.72
6 days		161.45	175.85	128.05	58.04	8.8
1 week		162.1	169.88	124.04	66.58	17.24
2 weeks		148.35	159.84	121.65	66.64	29.82
3 weeks		151.48	167.34	125.28	69.33	28.67
4 weeks		157.0	173.12	129.95	77.32	38.02
8 weeks		199.69	180.19	137.67	80.93	40.97

Table I-5 - CRPS for the Sheep River

	Date	22 June	23 June	24 June	25 June	26 June
Lead time						
1 day		353.04	149.19	33.95	12.84	21.05
2 days		858.77	362.04	8.02	10.18	20.08
3 days		888.16	419.65	106.88	5.65	19.02
4 days		887.51	425.49	133.34	25.13	22.14
5 days		882.2	425.12	137.34	41.39	51.22
6 days		881.47	419.15	135.89	44.49	66.26
1 week		880.43	419.6	131.09	42.27	69.37
2 weeks		864.63	410.33	126.71	33.78	60.87
3 weeks		852.62	387.73	96.12	13.03	25.91
4 weeks		864.44	407.51	117.09	29.26	47.53
8 weeks		857.1	404.42	107.7	21.01	33.4

CRPSS

Table I-6 - CRPSS at the Carseland dam [-]

	Date	22 June	23 June	24 June	25 June	26 June
Lead time						
1 day		0.81	0.77	0.76	0.87	0.89
2 days		-0.07	0.67	0.8	0.88	0.96
3 days		-0.14	-0.02	0.59	0.9	0.95
4 days		-0.05	-0.07	-0.01	0.46	0.93
5 days		0.01	-0.02	-0.04	0.02	0.67
6 days		0.03	0.01	-0.02	-0.02	0.13
1 week		0.04	0.03	0.01	-0.0	0.03
2 weeks		0.1	0.06	0.04	0.06	0.14
3 weeks		0.13	0.09	0.07	0.11	0.26
4 weeks		0.08	0.05	0.04	0.05	0.09
8 weeks		0.03	0.05	0.03	0.05	0.11

Table I-7 – CRPSS at Calgary for the Bow river [-]

	Date	**22 June**	**23 June**	**24 June**	**25 June**	**26 June**
Lead time						
1 day		0.68	0.79	0.92	0.95	0.87
2 days		0.01	0.63	0.93	0.95	0.96
3 days		-0.04	0.03	0.45	0.89	0.92
4 days		0.01	-0.0	0.09	0.8	0.8
5 days		0.04	0.01	0.05	0.3	0.97
6 days		0.05	0.04	0.06	0.18	0.38
1 week		0.06	0.05	0.09	0.2	0.24
2 weeks		0.08	0.08	0.12	0.27	0.3
3 weeks		0.09	0.08	0.13	0.29	0.37
4 weeks		0.06	0.06	0.09	0.19	0.21
8 weeks		0.02	0.05	0.08	0.19	0.21

Table I-8 - CRPSS at Calgary for the Elbow river [-]

	Date	**22 June**	**23 June**	**24 June**	**25 June**	**26 June**
Lead time						
1 day		0.77	0.67	0.74	0.76	0.76
2 days		-2.02	0.74	0.77	0.76	0.78
3 days		-2.11	-0.91	0.58	0.81	0.78
4 days		-1.54	-0.97	-0.53	0.53	0.84
5 days		-1.08	-0.75	-0.57	-0.41	0.5
6 days		-0.98	-0.52	-0.46	-0.44	-0.38
1 week		-0.81	-0.5	-0.35	-0.37	-0.41
2 weeks		-0.32	-0.34	-0.28	-0.25	-0.26
3 weeks		-0.43	-0.22	-0.12	-0.1	-0.1
4 weeks		-0.48	-0.27	-0.19	-0.17	-0.17
8 weeks		-0.09	-0.27	-0.2	-0.17	-0.17

Table I-9 – CRPSS at Lake Louise [-]

	Date	**22 June**	**23 June**	**24 June**	**25 June**	**26 June**
Lead time						
1 day		0.92	0.89	0.89	0.92	0.98
2 days		0.33	0.93	0.8	0.86	0.92
3 days		0.14	0.38	0.97	0.68	0.81
4 days		0.13	0.22	0.52	0.93	0.59
5 days		0.17	0.2	0.37	0.74	0.82
6 days		0.21	0.22	0.33	0.59	0.91
1 week		0.21	0.25	0.35	0.52	0.83
2 weeks		0.28	0.29	0.36	0.52	0.7
3 weeks		0.26	0.26	0.34	0.5	0.71
4 weeks		0.24	0.23	0.32	0.45	0.62
8 weeks		0.03	0.2	0.28	0.42	0.59

Table I-10 - CRPSS for the Sheep River [-]

	Date	**22 June**	**23 June**	**24 June**	**25 June**	**26 June**
Lead time						
1 day		0.59	0.63	0.69	0.42	0.43
2 days		0.01	0.11	0.93	0.54	0.45
3 days		-0.03	-0.04	0.02	0.74	0.48
4 days		-0.02	-0.05	-0.22	-0.14	0.4
5 days		-0.02	-0.05	-0.25	-0.88	-0.39
6 days		-0.02	-0.03	-0.24	-1.02	-0.8
1 week		-0.02	-0.04	-0.2	-0.92	-0.89
2 weeks		0.0	-0.01	-0.16	-0.54	-0.65
3 weeks		0.02	0.04	0.12	0.41	0.3
4 weeks		0.0	-0.01	-0.07	-0.33	-0.29
8 weeks		0.01	0.0	0.02	0.04	0.09

Appendix J – Dichotomous skill scores

22 June

Table J-1 - Dichotomous skill scores for the Bow at the 22nd of June

Lead time	Hit 1	False 1	Hit 2	False 2	Hit 3	False 3	Hit 4	False 4	Hit 5	False 5
1 day	1.0	-	1.0	-	1.0	-	0.708	-	0	0
2 days	1.0	-	1.0	-	0.0	-	0	-	0	0
3 days	1.0	-	1.0	-	0.0	-	0	-	0	0
4 days	1.0	-	1.0	-	0.01	-	0	-	0	0
5 days	1.0	-	1.0	-	0.02	-	0	-	0	0
6 days	1.0	-	1.0	-	0.02	-	0	-	0	0
1 week	1.0	-	0.92	-	0.02	-	0	-	0	0
2 weeks	1.0	-	1.0	-	0.06	-	0	-	0	0
3 weeks	1.0	-	0.92	-	0.08	-	0	-	0	0
4 weeks	1.0	-	0.83	-	0.07	-	0	-	0	0
8 weeks	1.0	-	0.58	-	0.08	-	0	-	0	0

Table J-2 - Dichotomous skill scores for the Elbow at the 22nd of June

Lead time	Hit 1	False 1	Hit 2	False 2	Hit 3	False 3	Hit 4	False 4	Hit 5	False 5
1 day	1.0	0.0	0.94	0.0	-	0	-	0	-	0
2 days	0.0	0.0	0.0	0.0	-	0	-	0	-	0
3 days	0.0	0.0	0.0	0.0	-	0	-	0	-	0
4 days	0.08	0.08	0.08	0.08	-	0	-	0	-	0
5 days	0.17	0.17	0.17	0.17	-	0	-	0	-	0
6 days	0.17	0.17	0.17	0.17	-	0	-	0	-	0
1 week	0.17	0.17	0.17	0.17	-	0	-	0	-	0
2 weeks	0.25	0.25	0.17	0.25	-	0	-	0	-	0
3 weeks	0.25	0.25	0.17	0.17	-	0	-	0	-	0
4 weeks	0.25	0.25	0.17	0.17	-	0	-	0	-	0
8 weeks	0.42	0.5	0.17	0.17	-	0	-	0	-	0

23 June

Table J-3 - Dichotomous skill scores for the Bow at the 23rd of June

Lead time	Hit 1	False 1	Hit 2	False 2	Hit 3	False 3	Hit 4	False 4	Hit 5	False 5
1 day	1.0	-	1.0	-	1.0	-	1.0	-	1.0	1.0
2 days	1.0	-	1.0	-	1.0	-	0.48	-	0.0	0.0
3 days	1.0	-	1.0	-	0.0	-	0.0	-	0.0	0.0
4 days	1.0	-	1.0	-	0.0	-	0.0	-	0.0	0.0
5 days	1.0	-	1.0	-	0.0	-	0.0	-	0.0	0.0
6 days	1.0	-	1.0	-	0.0	-	0.0	-	0.0	0.0
1 week	1.0	-	1.0	-	0.0	-	0.0	-	0.0	0.0
2 weeks	1.0	-	1.0	-	0.0	-	0.0	-	0.0	0.0
3 weeks	1.0	-	1.0	-	0.08	-	0.0	-	0.0	0.0
4 weeks	1.0	-	0.90	-	0.0	-	0.0	-	0.0	0.0
8 weeks	1.0	-	0.92	-	0.0	-	0.0	-	0.0	0.0

Table J-4 - Dichotomous skill scores for the Elbow at the 23rd of June

Lead time	Hit 1	False 1	Hit 2	False 2	Hit 3	False 3	Hit 4	False 4	Hit 5	False 5
1 day	1.0	-	1.0	-	-	0	-	0	-	0
2 days	1.0	-	1.0	-	-	0	-	0	-	0
3 days	0.0	-	0.0	-	-	0	-	0	-	0
4 days	0.0	-	0.0	-	-	0	-	0	-	0
5 days	0.08	-	0.08	-	-	0	-	0	-	0
6 days	0.17	-	0.17	-	-	0	-	0	-	0
1 week	0.17	-	0.17	-	-	0	-	0	-	0
2 weeks	0.22	-	0.17	-	-	0	-	0	-	0
3 weeks	0.20	-	0.17	-	-	0	-	0	-	0
4 weeks	0.21	-	0.17	-	-	0	-	0	-	0
8 weeks	0.17	-	0.14	-	-	0	-	0	-	0

24 June

Table J-5 - Dichotomous skill scores for the Bow at the 24th of June

Lead time	Hit 1	False 1	Hit 2	False 2	Hit 3	False 3	Hit 4	False 4	Hit 5	False 5
1 day	1.0	-	1.0	-	1.0	-	1.0	0.2	-	0.17
2 days	1.0	-	1.0	-	1.0	-	1.0	0.28	-	0.16
3 days	1.0	-	1.0	-	0.99	-	0.0	0.0	-	0.0
4 days	1.0	-	1.0	-	0.04	-	0.0	0.0	-	0.0
5 days	1.0	-	1.0	-	0.01	-	0.0	0.0	-	0.0
6 days	1.0	-	1.0	-	0.01	-	0.0	0.0	-	0.0
1 week	1.0	-	1.0	-	0.01	-	0.0	0.0	-	0.0
2 weeks	1.0	-	1.0	-	0.02	-	0.0	0.0	-	0.0
3 weeks	1.0	-	1.0	-	0.08	-	0.0	0.0	-	0.0
4 weeks	1.0	-	0.96	-	0.08	-	0.0	0.0	-	0.0
8 weeks	1.0	-	0.98	-	0.0	-	0.0	0.0	-	0.0

Table J-6 - Dichotomous skill scores for the Elbow at the 24th of June

Lead time	Hit 1	False 1	Hit 2	False 2	Hit 3	False 3	Hit 4	False 4	Hit 5	False 5
1 day	1.0	-	1.0	-	-	-	-	-	-	-
2 days	1.0	-	1.0	-	-	-	-	-	-	-
3 days	1.0	-	1.0	-	-	-	-	-	-	-
4 days	0.0	-	0.0	-	-	-	-	-	-	-
5 days	0.0	-	0.0	-	-	-	-	-	-	-
6 days	0.08	-	0.05	-	-	-	-	-	-	-
1 week	0.17	-	0.17	-	-	-	-	-	-	-
2 weeks	0.17	-	0.17	-	-	-	-	-	-	-
3 weeks	0.17	-	0.17	-	-	-	-	-	-	-
4 weeks	0.17	-	0.11	-	-	-	-	-	-	-
8 weeks	0.17	-	0.08	-	-	-	-	-	-	-

25 June

Table J-7 - Dichotomous skill scores for the Bow at the 25th of June

Lead time	Hit 1	False 1	Hit 2	False 2	Hit 3	False 3	Hit 4	False 4	Hit 5	False 5
1 day	1.0	-	1.0	-	1.0	0	-	0	-	0
2 days	1.0	-	1.0	-	1.0	0	-	0	-	0
3 days	1.0	-	1.0	-	1.0	0	-	0	-	0
4 days	1.0	-	1.0	-	0.99	0	-	0	-	0
5 days	1.0	-	1.0	-	0.08	0	-	0	-	0
6 days	1.0	-	1.0	-	0.08	0	-	0	-	0
1 week	1.0	-	1.0	-	0.08	0	-	0	-	0
2 weeks	1.0	-	1.0	-	0.08	0	-	0	-	0
3 weeks	1.0	-	1.0	-	0.08	0	-	0	-	0
4 weeks	1.0	-	1.0	-	0.07	0	-	0	-	0
8 weeks	1.0	-	1.0	-	0.01	0	-	0	-	0

Table J-8 - Dichotomous skill scores for the Elbow at the 25th of June

Lead time	Hit 1	False 1	Hit 2	False 2	Hit 3	False 3	Hit 4	False 4	Hit 5	False 5
1 day	1.0	-	1.0	-	-	0	-	0	-	0
2 days	1.0	-	1.0	-	-	0	-	0	-	0
3 days	1.0	-	1.0	-	-	0	-	0	-	0
4 days	1.0	-	1.0	-	-	0	-	0	-	0
5 days	0.0	-	0.0	-	-	0	-	0	-	0
6 days	0.0	-	0.0	-	-	0	-	0	-	0
1 week	0.08	-	0.0	-	-	0	-	0	-	0
2 weeks	0.17	-	0.07	-	-	0	-	0	-	0
3 weeks	0.17	-	0.11	-	-	0	-	0	-	0
4 weeks	0.15	-	0.08	-	-	0	-	0	-	0
8 weeks	0.10	-	0.08	-	-	0	-	0	-	0

26 June

Table J-9 - Dichotomous skill scores for the Bow at the 26th of June

Lead time	Hit 1	False 1	Hit 2	False 2	Hit 3	False 3	Hit 4	False 4	Hit 5	False 5
1 day	1.0	-	1.0	-	1.0	-	-	0	-	0
2 days	1.0	-	1.0	-	1.0	-	-	0	-	0
3 days	1.0	-	1.0	-	1.0	-	-	0	-	0
4 days	1.0	-	1.0	-	1.0	-	-	0	-	0
5 days	1.0	-	1.0	-	1.0	-	-	0	-	0
6 days	1.0	-	1.0	-	0.04	-	-	0	-	0
1 week	1.0	-	1.0	-	0.02	-	-	0	-	0
2 weeks	1.0	-	1.0	-	0.03	-	-	0	-	0
3 weeks	1.0	-	1.0	-	0.03	-	-	0	-	0
4 weeks	1.0	-	1.0	-	0.0	-	-	0	-	0
8 weeks	1.0	-	1.0	-	0.0	-	-	0	-	0

Table J-10 - Dichotomous skill scores for the Elbow at the 26th of June

Lead time	Hit 1	False 1	Hit 2	False 2	Hit 3	False 3	Hit 4	False 4	Hit 5	False 5
1 day	1.0	-	1.0	-	-	0	-	0	-	0
2 days	1.0	-	1.0	-	-	0	-	0	-	0
3 days	1.0	-	1.0	-	-	0	-	0	-	0
4 days	1.0	-	1.0	-	-	0	-	0	-	0
5 days	1.0	-	1.0	-	-	0	-	0	-	0
6 days	0.0	-	0.0	-	-	0	-	0	-	0
1 week	0.0	-	0.0	-	-	0	-	0	-	0
2 weeks	0.12	-	0.0	-	-	0	-	0	-	0
3 weeks	0.17	-	0.08	-	-	0	-	0	-	0
4 weeks	0.08	-	0.08	-	-	0	-	0	-	0
8 weeks	0.08	-	0.08	-	-	0	-	0	-	0