Multi-Scale Streamflow Simulation

by

Siavash POURYOUSEFI MARKHALI

THESIS PRESENTED TO ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
IN PARTIAL FULFILLMENT FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
Ph.D.

MONTREAL, 27 APRIL 2023

ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
UNIVERSITÉ DU QUÉBEC

ACKNOWLEDGEMENTS

Simulation multi-échelle du débit des cours d'eau

Siavash POURYOUSEFI MARKHALI

RÉSUMÉ

Les questions d'échelle représentent un problème non résolu en hydrologie. Les modèles hydrologiques distribués sont capables de prendre en compte l'hétérogénéité des bassins versants, mais on ne sait toujours pas dans quelle mesure les variations dans la représentation des échelles spatio-temporelles, dans ces modèles, entraînent une incertitude sur les simulations. De plus, la question de la valeur ajoutée d'une discrétisation spatio-temporelle plus fine doit être étudiée davantage. La présente thèse aborde ces sujets dans le contexte de la simulation des écoulements, de la projection des crues et de la régionalisation des paramètres des modèles hydrologiques.

Tous les bassins versants étudiés dans cette thèse sont situés dans la partie sud de la province de Québec, au Canada. Cette recherche repose sur l'utilisation de deux modèles hydrologiques distribués et à base physiques avec différents degrés de complexité (Hydrotel et WaSiM). Pour la simulation et la projection des inondations, les modèles sont calibrés en considérant quatre niveaux différents de spatialisation de l'information physiographique et des paramètres des modèles hydrologiques. Les données du projet ClimEx sont corrigées pour des pas de temps de 3 et 24 heures en utilisant la méthode de correction de biais multivariée n-dimensionnelle (MBCn), et utilisées comme entrées dans les modèles hydrologiques pour projeter le débit sur la période 1991-2100.

Les résultats montrent que la variation de la résolution temporelle n'a que des impacts mineurs sur l'incertitude des simulations en conditions historiques, et que son impact dépend du choix du modèle hydrologique. Le modèle le plus sophistiqué (WaSim) présente une plus grande incertitude. Quant à la variation de la discrétisation spatiale, elle peut entraîner des incertitudes pour les bassins versants à faible pente ou comportant des zones accidentées.

En ce qui concerne la projection des crues, en affinant l'échelle temporelle, les résultats montrent que la fréquence et l'amplitude des débits extrêmes en été augmentent dans le futur. De plus, le choix du modèle hydrologique pour la projection des débits de crues est plus important pour les grands bassins versants. Enfin, il n'existe pas de patron caractéristique concernant l'incertitude liée à la résolution spatiale et à la taille du bassin versant. Cependant, cela affecte la direction et le niveau de signifitaction des tendances observées pour les débits extrêmes dans les simulations.

Cette thèse propose et teste également une méthode de régionalisation basée sur les forêts aléatoires (RF). Elle est appliquée aux paramètres d'Hydrotel à différentes échelles spatio-temporelles. Les résultats montrent que la technique de régionalisation proposée est plus performante pour des pas de temps plus courts. De plus, les paramètres régionalisés sont spatialement cohérents. En fin de compte, l'utilisation de descripteurs des bassins versant ayant une meilleure représentativité spatiale entraîne une amélioration (de plus de 10%) pour les simulations avec un pas de temps de 24 heures.

VIII

Multi-Scale Streamflow Simulation

Siavash POURYOUSEFI MARKHALI

ABSTRACT

Scale issues represent an unsolved problem in hydrological sciences. Distributed hydrological models are capable of accounting for catchment heterogeneity, but it remains unclear to what extent the variations in the representation of spatio-temporal resolution in these models leads to uncertainty in the simulations. Moreover, the added value of more refined spatio-temporal discretization is also unclear. The present thesis addresses these topics in the context of streamflow simulation, flood projection and regionalization of model parameters .

All catchments studied in this thesis are located in Southern Quebec, Canada . This research uses two process-based distributed hydrological models with different degrees of complexity (Hydrotel and WaSiM). For the flood simulation and projection, the models are calibrated with four different levels of spatial discretization of physiographic data and in models' parameters. The climate extreme project (ClimEx) dataset is bias corrected for 3- and 24-hour time-steps using the n-dimensional multivariate bias correction method (MBCn), and used as inputs to the hydrological models to project streamflow over the 1991-2100 period.

The results show that the variation of temporal resolution has only minor impacts on the uncertainty of historical simulations, and the impact depends on the choice of the model. The more sophisticated model (WaSim) has a larger uncertainty. As for varying the spatial discretization, it can cause uncertainties for catchments with low slopes or uneven areas.

Regarding flood projection, by refining the temporal scale, the results show that both the frequency and amplitude of extreme summer-fall flow increases in the future. Moreover, the choice of hydrological model for flood projection is more important for larger catchments. Finally, no distinct pattern exists regarding the uncertainty related to the spatial resolution and catchment size. However, this affects the direction and significance of the trends observed for extreme flow in the simulations.

This thesis also proposes and tests a regionalization method based on random forests (RF). It is applied to the parameters of Hydrotel at different spatio-temporal resolutions. The results show that the proposed regionalization technique performs better for shorter time-steps. Moreover, the regionalized parameters are spatially consistent. In the end, using catchment descriptors that have a better spatial representativity results in an improvement (more than 10%) in the simulations using a 24h time-step.

Keywords: distributed models, spatio-temporal discretization, flood simulation, regionalization

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Page

LIST OF ABBREVIATIONS

AET	Actual Evapotranspiration
ANOVA	Analyse of Variance
ANN	Artificial Neural Network
AOGCMs	Atmosphere-Ocean General Circulation Models
CAD	Canadian Dollar
CD	Catchment Descriptor
CanESM2	Canadian Earth System Model version 2
ClimEx	Climate Extremes Project
CRCM5	Canadian Regional Climate Model version 5
CP	Calibration Parameters
DDS	Dynamically Dimensional Search
DEH	Direction de l'Expertise Hydrique
DEM	Digital Elevation Model
DL	Deep Learning
ECMWF	European Centre for Medium-Range Weather Forecast
ECDF	Empirical Cumulative Distribution Function
ESMs	Earth System Models
ERA5	ECMWF ReAnalysis5
GCMs	General Circulation Models

GEV	Generalized Extreme Value distribution
GHM	Global hydrological model
HRU	Hydrological Response Unit
IAHS	International Association of Hydrological Sciences
KGE	Kling-Gupta Efficiency
LSM	Land Surface Model
LSTM	Long Short-Term Memory
MBC	Multivariate Bias Correction
MELCCC	Ministère de l'Environnement et de la Lutte
ML	Machine Learning
MOPEX	Model Parameter Estimation Experimen
MPR	Multiscale Parameter Regionalization
NSE	Nash-Sutcliffe Efficiency
PUB	prediction in ungauged basins
QM	Quantile Mapping
RCM	Regional Climate Model
RCP	Representative Concentration Pathway
RF	Random Forest
RHHU	Relatively Homogeneous Hydrological Units
RNN	Recursive Neural Network

SD   Snow Depth

WaSiM   Water Simulation Model

LIST OF SYMBOLS AND UNITS OF MEASUREMENTS

°C	degree Celsius
m	meters
km	kilometre
s	second
h	hour

INTRODUCTION

Flooding is the most frequent natural disaster in the world. It is estimated that 43 % of total natural disasters and 47 % of total climatic disasters are related to floods (Kundzewicz *et al.*, 2019a). In terms of humanitarian costs, the fatality of river floods in the last century amounts to 7 million people, and since 1985, 2.3 billion people, directly or indirectly, have been affected by floods (Gaur, Gaur & Simonovic, 2018). In terms of economic losses, the annual average costs directly related to flooding events are estimated to US $ 104 billion (for Disaster Reduction. Secretariat, 2015). The cost of floods due to the rise of temperature by 1.5° relative to the baseline of 1976-2005 are projected to significantly increase with the order of 2 and 3 for human and economic losses respectively (Dottori *et al.*, 2018).

Three major interconnected components and their related processes define flood risk, namely flood hazard, exposure, and vulnerability (Barendrecht, Viglione & Blöschl, 2017; Vorogushyn *et al.*, 2018; Di Baldassarre *et al.*, 2015). Flood hazard is the hydro-climate processes causing floods. Exposure is the elements, including the human population and infrastructures, that are at risk of being impacted by floods. Vulnerability is the degree to which the components at risk are sensitive to floods (for Disaster Reduction. Secretariat, 2015). Historical records have shown that flood damage has been increasing in recent decades because of land use change such as urbanization, and the population growth. Such developments resulted in elevated human and infrastructure exposure to floods and, consequently, rising vulnerability. Databases of natural disasters (e.g. Floods & Floods, 2004) show increasing economic costs due to floods across different spatial scales. Expressed by a percentage of Gross Domestic Product (GDP), the annual economic losses attributed to hydrological risks in terms of hazard and exposure is around 0.1 to 1 % of GDP depending on the level of development of countries (Field, Barros, Stocker & Dahe, 2012; Kreibich *et al.*, 2017). However, yearly exposure and costs associated to flood hazards in some countries reach more than 10 percent of the GDP (e.g., India, Bangladesh, Vietnam,

and Cambodia). In absolute terms, US and China spend more than 10 billion USD per year to confront the losses inflicted by floods (Kundzewicz *et al.*, 2014).

Canada, on average, spends one to 2 billion CAD per year to cope with flood risk and its consequences. This makes flooding the costliest natural disaster throughout the country(Oubennaceur, Chokmani, Nastev, Lhissou & El Alem, 2019). In the past decade, Canada has been hit by multiple floods, particularly in urban areas (e.g., Montreal in 2012, Calgary in 2013, Toronto in 2005, 2013, and British Columbia in 2021). Over the recent decade, Quebec has witnessed multiple floods due to a combination of snowmelt and rainfall (Buttle *et al.*, 2016). Vast areas of the province were flooded, causing damage to infrastructures, isolation of properties, and evacuation of thousands of residents.

Analyzing some key streamflow indices shows an increasing trend of streamflow across northern regions and a decreasing trend across southern regions of Canada (Burn & Elnur, 2002). In addition, the long-term projection of streamflow in the future shows a change in timing of the peak, magnitude, and frequency of floods across different regions of Canada (Leduc *et al.*, 2019; Martel, 2019). These results suggest that the pattern of flooding is changing in Nordic catchments. These catchments are prone to two types of flood events: spring floods due to snowmelt and summer-fall floods associated with intense short-duration rainfalls. Climate change is expected to attenuate spring floods and amplify summer-fall floods in magnitude and frequency (Donat, Lowry, Alexander, O'Gorman & Maher, 2016). Studies show that the rise of temperature increases the probability of intense short-duration rainfall at sub-daily time-scales, which can translate into more extreme flows and potential floods for small to medium catchments, with a relatively short concentration time (Westra *et al.*, 2014). Consequently, it is possible that these catchments will become more vulnerable to summer-fall floods.

Future flood risks are typically evaluated using a hydroclimatic modeling chain, which includes global climate models (GCMs), downscaling and/or post-processing of climate model outputs,

and hydrological models. GCMs project climate based on a series of Shared Socio Economic Pathways(SSPs) . Their outputs require downscaling and/or post-processing to increase the resolution of/bias correct the projections of the different climate variables. Hydrological models are used to simulate streamflow and other hydrological variables(Merz *et al.*, 2021). Each step contributes to the total uncertainty, propagating top-down through the chain (Chen *et al.*, 2019b). Several studies have quantified the uncertainties corresponding with each step of this process (e.g., Chen, Brissette, Poulin & Leconte, 2011; Poulin, Brissette, Leconte, Arsenault & Malo, 2011; Meresa & Romanowicz, 2017; Meresa *et al.*, 2022). However, less attention has been given to the uncertainty linked to representing the heterogeneity of the catchments with more or less details. This is mainly because flood generation processes have not yet been completely understood (Blöschl & Sivapalan, 1995). Non-linear and complex catchment responses across different spatio-temporal resolution and lack of a scale-relevant theory complicate modeling practices (Beven, 2010; Blöschl *et al.*, 2013). The dearth of knowledge and the complexity of hydrological processes further lead to a parameterization that is not scale-invariant (i.e. sensitive to scale of modeling) with the parameters mostly lacking tangible physical meaning (Samaniego, Kumar & Attinger, 2010; Samaniego *et al.*, 2017).

The advance of computational power facilitates the use of sophisticated distributed process-based Hydrological models to study fluxes and state variables in fine details across catchments (Sidle, 2021). This, however adds to the complexity of streamflow and flood simulations. Since distributed models need information about the physiographic characteristics of the catchments, the representation of catchment heterogeneity is an essential factor. The decision to represent spatial heterogeneity in distributed models depends on the spatial discretization. The choice should also be considered in conjunction with the temporal resolution. Regarding continuum-based models, for example, a finer temporal resolution might demand higher spatial resolution to satisfy the conditions required for computationally stable responses (Thober *et al.*, 2019). In addition, the catchment's key topographic and physiographic features may substantially influence

the hydrological short-term response. For instance, urbanization will increase imperviousness and create complicated routing pathways (Cao *et al.*, 2020a). The tendency to use coarser grids to compensate for computational costs can eliminate important underlying information, alter model parameterization, and produce different model responses, further increasing the uncertainty of hydrological simulations. This issue should be seen in the context of climate change impact assessment, where characterizing uncertainties associated with different modeling stages is necessary to develop coherent policies and devise effective adaptation strategies (Kundzewicz *et al.*, 2019a).

The present thesis is an effort to understand the importance of spatio-temporal resolution in simulating streamflow and floods with process-based distributed Hydrological models. This thesis focuses on the three following topics which define the research objectives: (1) quantifying the uncertainties linked to spatial heterogeneity of catchments in streamflow and flood simulations, (2) performing multi-scale flood projections under future climate, and (3) evaluating the effect of spatio-temporal resolution on streamflow regionalization. Multiple catchments with diverse sizes and hydro-climatic regimes in the province of Quebec have been used to address these topics. This thesis is a part of a Quebec-Bavarian collaboration to simulate climate extremes (ClimEx).

The structure of the thesis is as follows. Chapter 1 provides a literature review associated with the research problems and objectives. Chapter 2 summarizes the materials and methods. Chapters 3 to 5 are formed of three journal articles addressing the three aforementioned research objectives. Chapter 6 provides a general discussion to clarify the results. A conclusion highlights the main findings and provides recommendations for future studies.

CHAPTER 1

LITERATURE REVIEW

This chapter presents a review of the existing literature relevant to the objectives of this thesis. Section 1.1 gives a general review of flood generation mechanisms. Section 1.2 reviews the components of climate change impact assessment modeling chain. Section 1.3 presents a general overview of hydrological. Section 1.4 details the uncertainties involved in hydro-climatic projections. Section 1.5 highlights the importance of choosing the appropriate spatio-temporal scale in the context of hydrological. Finally, section 1.6 highlights the importance of scale in the regionalization of hydrological models.

1.1 River Flood Generation

The Intergovernmental Panel of Climate Change Special Report of Extreme (IPCC SREX Field *et al.*, 2012) classifies floods into various types, including "river (fluvial) flood, flash flood, urban flood, pluvial flood, coastal flood,", for which the climate forcing, mechanisms and governing processes are different. Multiple climatic drivers can lead to river flood generation: prolonged precipitation, intense rainfall, monsoon, snow melt, rain on snow, etc. The distinct interplay of location, time of the year, and catchment condition are referred to as a flood regime. Spring freshet, for example, is typical of Northern Europe and Quebec (Riboust & Brissette, 2016; Blöschl *et al.*, 2017; Sivapalan, Blöschl, Merz & Gutknecht, 2005). The principal driver of flood events in Australia is intense short-duration rainfall (Ishak & Rahman, 2019). For Northwestern Europe, prolonged rainfall in the winter and high soil moisture are the main flood drivers (Bertola, Viglione, Lun, Hall & Blöschl, 2020; Kemter, Merz, Marwan, Vorogushyn & Blöschl, 2020). Extreme floods however contain surprising elements not necessarily related to a specific flood regime; examples include: Typhoon tracks in Thailand 2011 (Petvirojchai & SaraPa, 2018), atmospheric blocking in 1993 in Mississippi and in the 2011 Pakistan floods (Francis & Vavrus, 2012; Hong, Hsu, Lin & Chiu, 2011; Grams, Binder, Pfahl, Piaget & Wernli, 2014), as well as atmospheric rivers that triggered extreme floods in various parts of the world (Merz *et al.*, 2021).

While climate drivers are the most important cause of floods, antecedent catchment conditions can modulate the severity of the events. When catchment wetness is high, even a small amount of precipitation can transform into direct surface runoff and generate flooding (Nakamura, Lall, Kushnir, Robertson & Seager, 2013). Such a combination was observed in 2013 in central Europe, when atmospheric anomalies generated prolonged precipitation that elevated soil moisture. This provided favorable conditions for flooding with a modest rainfall (Nakamura *et al.*, 2013).

Physiographic characteristics also contribute to catchment response to climate drivers. Land use change can influence catchment behavior for medium to small catchments. For instance, catchments with a large proportion of agricultural lands have compact soil as a result of heavy agricultural machinery operations. This leads to lower infiltration rate and direct contribution of precipitation into flood (Keller, Sandin, Colombi, Horn & Or, 2019). In urban catchments, in addition to low infiltration rate, the sewage network makes hydrological pathways shorter and increases flow peaks (Miller & Hutchins, 2017). Deforestation also changes macropores' structure, reduces permeability, and alters preferential flow pathways, leading to increasing flow peaks (Gao *et al.*, 2018). Land use change has only minor effects on large catchments, as studies have shown that it may contribute to only a few more centimeters of peak water depth with relatively smaller effect on large territories(Bronstert *et al.*, 2007; Te Linde, Aerts & Kwadijk, 2010). Moreover, land use can merely affect extreme floods in the case of near saturated soil, where land cover has insignificant effects on flood generation (Rogger *et al.*, 2017). However, the relationship between flood and land use at different scales is complex and has not been studied extensively.

Compared to land use, the impact of hydraulic structures and human intervention on peak flow is better understood. In most cases, simple mass and momentum equations can be used to simulate flood peaks (Horváth *et al.*, 2020; Buttinger-Kreuzhuber, Horváth, Noelle, Blöschl & Waser, 2019). Dams typically reduce flood hazard by storing water in reservoirs (Volpi, Di Lazzaro, Bertola, Viglione & Fiori, 2018; Mei, Van Gelder, Dai & Tang, 2017). For floods with large return periods, the capacity to dampen the flood wave declines (Vorogushyn, Lindenschmidt,

Kreibich, Apel & Merz, 2012; Volpi *et al.*, 2018). Channel straightening affects the peak flow in different ways. In local reaches, increasing channel capacity results in less frequent floodplain inundation but has minimal effects on the peak. In downstream reaches, since the flood plain storage has been removed, the limited dampening effect increases flood hazard (Blazejewski, Pilarczyk & Przedwojski, 1995). This also depends on the flood magnitude. For small to medium floods, channel straightening increases flood hazard downstream because the peak is not modulated by the flood plain early on. However, for extreme floods, the activated floodplain helps dampening flood wave (Blöschl, 2022b). Severe floods may cause damages to hydraulic structures (i.e., dams, levees, etc.) and create unexpected events resulting in catastrophic costs (Merz *et al.*, 2021).

Human influence on streamflow is not limited to engineering works. Anthropogenic climate change affects precipitation patterns, temperature, snow melt, etc. A recent meta-analysis by (Merz *et al.*, 2021) shows a positive future trend for flood hazards in many parts of the world, including North America. Assessing future changes requires a modeling chain, starting from SSPs, which are used as input for climate models. The output of climate models are then used to force hydrological models (Hall *et al.*, 2014). Climate projections can be obtained at different scales, which helps to conduct impact assessment studies locally, regionally, or globally (Do *et al.*, 2020; Swain *et al.*, 2020). The accuracy of these studies directly depends on how climate models are capable of representing atmospheric processes at different scales (Blöschl, 2022b). Several studies have shown that warming temperatures can result in extreme precipitation and higher flood risks in the future. Similar to land use impact, the impact of climate change should be studied separately for small and large catchments, because small catchments are more sensitive to intense short-duration rainfall than large catchments (Blöschl, 2022a).

1.2   Climate Models

Global Climate Models (GCMs) consist of a dynamical core for solving the primitive equations governing the atmosphere (e.g., mass, momentum, energy (Bjerknes, 1910)) across time and space. The Coupled Model Intercomparison Project (CMIP) provides the outputs of several

GCMs, for use in hydrological (or other) studies. The project includes CMIP3 (Meehl *et al.*, 2007), CMIP5 (Taylor, Stouffer & Meehl, 2012), and CMIP6 (O'Neill *et al.*, 2016) phases, for different scenarios (Gurney *et al.*, 2022). CMIP3 was based on the generation of Special Report on Emission Scenarios (SRES; Meehl *et al.*, 2007), while the most recent CMIP5 and CMIP6 experiments are based on Representative Concentration Pathways (RCPs). RCPs are socio-economic pathway scenarios designed to represent anthropogonic climate forcings into climate models. These scenarios are based on greenhouse gas concentration in the atmosphere and expressed by the total radiative forcing (in W/m^2) by the year of 2100 (IPCC, 2014). The main RCPs used for running the climate models are RCP2.6 (very low forcing), RCP4.5 (medium forcing), RCP6 (medium forcing), and RCP8.5 (very high forcing) (Van Vuuren *et al.*, 2011). Apart from RCPs, the Shared Socio-economic Pathways (SSPs) have recently been developed for CMIP6 (Masson-Delmotte *et al.*, 2021). These scenarios consider the socio-economic developments, including Gross Domestic Product (GDP) growth, the rise of population, education, Green House Gas (GHG) emission, technological development, etc. These scenarios work in combination with RCPs to address the challenges regarding adaptation and mitigation strategies (Wei *et al.*, 2018).

Since the 1980s, climate models have moved towards the inclusion of more Earth processes and sub-processes into a single framework called Earth System Models (ESMs Hill, DeLuca, Suarez, Da Silva *et al.*, 2004). These models are the product of coupled Atmospheric and Ocean GCM (AOGCM) with other Earth system-related models such as Land-Surface Models (LSMs) and Global Hydrological Models (GHMs). GCMs are computationally demanding. To overcome this issue, the models are implemented on relatively coarse spatial grids (typically 100km) and temporal resolution (typically daily)(Navarro-Racines, Tarapues, Thornton, Jarvis & Ramirez-Villegas, 2020; Chen *et al.*, 2011). As a result, the outputs of GCMs are not suitable for experiments at regional or local scales and/or at sub-daily time scales (Chen *et al.*, 2011). This issue is particularly important for hydrologists, as most hydrological variables should be studied at scales that are finer than those of GCMs (Szolgayova, Laaha, Blöschl & Bucher, 2014; Zhang,

Viglione & Blöschl, 2022). This is typically resolved by downscaling the output of climate models, which is discussed in the next sub-section.

1.2.1 Downscaling

Two major approaches have been proposed for downscaling: statistical and dynamical.

In dynamical downscaling, the GCM's simulation is mapped to a specific region employing the output of the latter model (e.g., winds, pressure, temperature, humidity etc.) as boundary conditions for a high-resolution Regional Climate Model (RCM; Dickinson, Errico, Giorgi & Bates, 1989). RCMs are complete climate models but they have a higher topographic resolution, and often more detailed land/water interfaces, land use/land cover than GCMs (Buonomo, Jones, Huntingford & Hannaford, 2007; Vautard *et al.*, 2021). The main advantage of dynamic downscaling is that RCMs create meteorologically consistent climate variables and account for complex climatic processes and their non-linear interactions (Williams, Erickson & Petrzelka, 2010). Yet RCMs are susceptible to uncertainty and biases such as overestimation of drizzling and underestimation of extreme convective rainfall (Chen, Brissette, Chaumont & Braun, 2013; Bresson, Laprise, Paquin, Thériault & de Elía, 2017; Maraun *et al.*, 2010; Solman & Blázquez, 2019). It is a common practice to bias correct the output of RCMs (Faghih, Brissette & Sabeti, 2022)

Statistical downscaling seeks the relationship that maps large-scale climate simulation provided by GCMs to observed local climate for a reference period, then projects it into the future (Wood, Leung, Sridhar & Lettenmaier, 2004). The outputs of the GCM serve as a predictors for the local climate (Martel, Mailhot, Brissette & Caya, 2018). Statistical downscaling methods can be further divided into three categories: regression-based methods, weather typing, and stochastic modeling (Trzaska & Schnarr, 2014). For regressions, a transfer function between GCM (predictor) and observed data (predictant) is established. The choice of predictors and transfer function is generally arbitrary and based on expert knowledge and process understanding. Weather typing classifies similar patterns in the reference period for predictors, establishes

a relationship between the predictor and predictant in the reference period, and applies it to the future (Van Uytven, De Niel & Willems, 2020). In stochastic modeling, the parametric distribution of the predictant is conditioned to the distribution of the predictors. The most popular form of this method is the weather generator (Wilks, 1992; Chen, Chen & Guo, 2018).

1.2.2   Post-processing

The outputs of RCMs are susceptible to biases (Chen *et al.*, 2019a). Here we briefly review the most common bias correction approaches.

The Delta Approach consists in perturbing the observed precipitation and temperature using appropriate coefficients (Teutschbein & Seibert, 2013). The perturbation coefficient for temperature is the difference between the expected value for the future period and that of the reference period in the GCM or RCM domain. For precipitation, a multiplicative coefficient is used. This coefficient is the ratio of the expected value of the future period to the reference period. The main advantage of this approach is its simplicity. The main disadvantage is that the shape and type of the projected probability distribution remain the same as the historical one (Willems, Ntegeka, Baguis & Roulin, 2010).

Quantile Mapping (QM) is a robust post-processing approach that compares the distribution of historical and future variables. The differences between the quantiles of the two distributions is used to correct the projections (Willems *et al.*, 2010). In the case of precipitation, for example, an empirical distribution is usually preferred, considering both wet days and dry days in its construction (Chen *et al.*, 2013). QM corrects both the bias and dispersion of the projected distributions.

1.3   Hydrological modeling

Hydrological models have been classified into different categories based on their level of spatial discretization (lumped vs. distributed), and description of physical processes (conceptual vs. physical).

Lumped models consider the catchment as homogeneous. These models typically solve simple mass balance equations represented by parametric functions (Beven, 2010). They are parsimonious, and usually the data related to the physiography of the catchment is not required (Hrachowitz & Clark, 2017). Lumped models require minimal computational power, but their application is limited to cases where the end result (streamflow at the outlet) is what matters. HBV (Bergström *et al.*, 1995), SUPERFLEX (Fenicia, Kavetski & Savenije, 2011) and GR4J (Perrin, Michel & Andréassian, 2003) are examples of popular lumped models, even though they can also be spatialized. This has been the case for HBV, for instance (Lindström, Johansson, Persson, Gardelin & Bergström, 1997).

Distributed models account for the spatial heterogeneity of the catchments (Beven, 2011), with more or less detailed spatial discretization. For instance, the level of spatial scale can be managed by SWAT (Arnold, Srinivasan, Muttiah & Williams, 1998) and VIC (Liang, Lettenmaier, Wood & Burges, 1994) but is limited to sub-catchment (i.e. semi-distributed) whereas MIKE-SHE (Refsgaard & Storm, 1990) and HYDRUS-3D (Šimunek, van Genuchten & Šejna, 2008) are grid-based, fully distributed. There exist mainly three approaches for spatial discretization: 1) the discretization of input data with fixed parameters; 2) the discretization of parameters only; and 3) a combination of both approaches (Ajami, Gupta, Wagener & Sorooshian, 2004; Das, Bárdossy, Zehe & He, 2008; Euser, Hrachowitz, Winsemius & Savenije, 2015). Distributed models can be data and computationally-intensive, with a finer spatial resolution leading to an increase in computational time (Bierkens, 2015).

Physics-based models attempt to solve the governing equations of hydrological processes and describe fluxes and states and their complex relationships in heterogeneous catchment media (Clark *et al.*, 2017; Fatichi *et al.*, 2016). Theoretically, physics-based models should not need calibration, as their parameters should represent physical processes and catchment characteristics. However, this is not truly the case and calibration is still needed. Scale issues and catchment heterogeneities (e.g., soil, vegetation) are some of the factors that contribute to parameter uncertainty in physics-based modeling. (Wagener, Wheater & Gupta, 2004). As a result, identifying the parameters requires calibration of the model, which is data intensive

and computationally demanding. MIKE-SHE (Refsgaard & Storm, 1990) and HYDRUS-3D (Šimunek *et al.*, 2008) are among this type of models.

Conceptual (or bucket-type) models represent hydrological processes through macro-scale conceptualization (Hrachowitz *et al.*, 2013) with a simple parametrization. In some cases, the number of processes that these models take into account is limited (Clark *et al.*, 2015). These models are often but not exclusively implemented at the catchment or sub-catchment scales. They are typically neither computationally demanding nor data intensive. HBV (Bergström *et al.*, 1995), GR4J (Perrin *et al.*, 2003), mHM (Samaniego *et al.*, 2010) and CEQEAU (St-Hilaire *et al.*, 2015b) belong to this category.

A typical procedure for implementing a hydrological model after selecting the structure is to estimate parameter values using a systemic calibration against a portion of the observed data (Wagener *et al.*, 2001). The calibration procedure is based on finding a parameter set for which the simulated streamflows corroborate as much as possible observed streamflows. This similarity is measured by a metric, for instance the Nash–Sutcliffe efficiency (NSE Nash & Sutcliffe, 1970) or the Kling Gupta efficiency (KGE; Gupta, Kling, Yilmaz & Martinez, 2009). There exist many other metrics, and they can be selected according to the application. Calibration is an iterative process that involves trying multiple parameter sets (Ajami *et al.*, 2004). There exist various optimization algorithms and goodness of fit measures in the literature. Dynamically Dimensioned Search (DDS; Tolson & Shoemaker, 2007) and Shuffled Complex Evolution Algorithm (SCEA; Duan, Gupta & Sorooshian, 1993) are amongst the most popular search algorithms in hydrological modeling.

There often exist more than one acceptable solution (parameter set). This is referred to as equifinality (Beven, 2006). Therefore, in practice, we are dealing with a range for each parameter leading to parameter bounds instead of a unique answer (Wagener & Wheater, 2006). For impact assessment, it is essential to report the uncertainty related to model parameters for a realistic impact assessment .

1.4   Uncertainty

Beven (2016) divided uncertainty into two broad categories: epistemic and aleatory uncertainty. The former is generally due to a lack of understanding and can be reduced by acquiring new knowledge. The latter, however, is owing to "the randomness of a phenomenon," which is not reducible. The definition provided by the Intergovernmental Panel on Climate Change (Pachauri *et al.*, 2014) for uncertainty is "complete absence of information or presence of partial knowledge or information or lack of consensus over known or knowable phenomenon." A lack of knowledge or imperfect knowledge about the possible outcomes (e.g. climate change) leads to a less solid basis for treating that phenomenon with probability theories (Foley, 2010). The phenomenon, therefore, has to be studied by developing a series of plausible scenarios and accounting for corresponding uncertainties.

The Earth's climate is a complex system affected by a range of external forces (e.g., greenhouse effect, solar radiation, celestial body collision, land-surface, volcanic eruption, etc.) along with its internal variation (e.g., El Niño Southern Oscillations (ENSO), North Atlantic Oscillation (NAO), Arctic Oscillation (AO), Atlantic Multidecadal Oscillation (AMO), Pacific Decadal Oscillation (PDO), etc.). Such high complexity and degree of freedom are difficult to be represented by climate models. In addition, the dynamic interaction between various processes along with their feedback loops is yet to be studied and further reflected in models to adequately represent the climate system (Kundzewicz & Stakhiv, 2010). Given current knowledge and computational abilities, climate models are susceptible to a high degree of uncertainty.

The cascade of uncertainties in hydroclimatic modeling can be summarized as follows: 1) future socio-economic development, 2) selection of greenhouse gas emission scenarios, 3) global climate models (GCMs), 4) dynamical/statistical downscaling, 5) bias correction method, 6) observational uncertainty of hydroclimatological data, 7) structural uncertainty of hydrological model, and 8) parameterization of the hydrological model. The top-down approach of such a modeling chain guarantees uncertainty propagation through the models as every stage inherits the uncertainty introduced from the previous one.

The chain of uncertainty begins with the uncertainty of future developments in terms of economy and technology as well as population growth. As a result, prescribed scenarios for GHG emissions have been proposed, which are based on two pioneering approaches: the first one proposed by IPCC Special Report on Emission Scenarios – SRES (IPCC) and the second one based on Representative Concentration Pathways (Meinshausen *et al.*, 2011). Adopting each scenario for simulation may produce an entirely different output leading to considerable uncertainty, particularly for regional and local scales. Uncertainty of GHG scenarios further depends on the horizon of simulations. The effect of scenarios on the short-term is smaller compared to the distant future, as climate response in the short-term is heavily dependent on historical GHG concentrations (Meinshausen *et al.*, 2011).

Several GCMs have been developed and implemented, providing several projections for climate statistics (e.g., Taylor *et al.*, 2012). Different GCMs running with the same GHG scenarios can provide divergent projections. Different models represent physical sub-processes differently, including numerical methods and parameterizations (Chen *et al.*, 2011; Chen, Brissette, Lucas-Picher & Caya, 2017). Previous research results show that the contribution of inter-model uncertainty outweighs other sources of uncertainty in hydroclimatic modeling (Minville, Brissette & Leconte, 2008; Wilby & Harris, 2006).

Though GCMs are often recognized as the major source of uncertainty in impact assessment studies, hydrological modeling uncertainty also needs to be considered (Chen *et al.*, 2011). Hydrological uncertainty stems from data (scarcity, measurement error, scale mismatch, aggregation/disaggregation), the choice of hydrological model, and its parameterization (Kundzewicz *et al.*, 2018). Perhaps the most well-known hydrological uncertainty assessment framework is the generalized likelihood uncertainty estimation (GLUE) (Beven & Binley, 1992; Setegn, Srinivasan & Dargahi, 2008; Blasone *et al.*, 2008; Uhlenbrook, Seibert, Leibundgut & Rodhe, 1999; Wechsler, 2007). The framework divides the whole possible parametric space into behavioral and non-behavioral solutions by applying the condition of maximum likelihood of simulation to observations. The choice of maximum likelihood function and the threshold to

distinguish between behavioral and non-behavioral simulations are subjective, which has ensued controversies regarding the adequacy of the framework (Clark, Kavetski & Fenicia, 2011)

In hydrology, the uncertainty linked to model structure is referred to as structural uncertainty (Krysanova *et al.*, 2018). Each hydrological model is a combination of several hypotheses and algorithms (Euser *et al.*, 2013; Clark *et al.*, 2011). The choice of process description, numerical method, spatial and temporal scales, as well as all choices involved in the classification and representation of topography, soil type, vegetation, etc., are all included in structural uncertainty (Krysanova *et al.*, 2018). A multimodel approach allows for a simple and effective way of characterizing structural uncertainty (Thiboult, Anctil & Boucher, 2016; Butts, Payne, Kristensen & Madsen, 2004; Seiller, Anctil & Roy, 2017). The idea is simply to use an ensemble of models, ideally very different in their representation of hydrological processes (Krueger *et al.*, 2010; Buytaert & Beven, 2011).

1.5 Modeling Scales

1.5.1 Temporal Scale

Theoretically, extreme rainfall depends on surface temperature. By increasing the temperature, the Clausius-Clapeyron (CC) equation (Ivancic & Shaw, 2016) predicts that the capacity of the atmosphere to accumulate moisture also increases (Kharin, Zwiers, Zhang & Hegerl, 2007; Trenberth, Covey & Dai, 2018; Westra *et al.*, 2014). The theory suggests that per $1°C$ increase in temperature, the saturation specific humidity of the air will increase by 7% , resulting in higher available moisture for extreme precipitation (Pall, Allen & Stone, 2007). In practice, the scaling is not linear (O'Gorman, 2015). Using observed available data from 1929-2017, Yin *et al.* (2018) showed that three major consequences are associated to rising temperature: first, a monotonic increase of extreme precipitation; second, a monotonic decrease of extreme precipitation; and third, a hook-like behavior (Berg & Sheffield, 2018; Drobinski *et al.*, 2020) with increasing extreme precipitation up to a certain threshold of temperature and a decrease afterward. Higher latitudes, including most of Canada, have either witnessed a monotonic increase of extreme

precipitation corresponding with a rise in temperature or a hook-like structure with a threshold temperature before which the extreme precipitation has been increased (Yin *et al.*, 2018). Their analyses regarding high flow further confirm that the hook-like structure is a dominant pattern worldwide except for the tropics.

Rising temperature can further change the dominant pattern of rainfall from stratiform to convective (Berg & Haerter, 2013). Since convective rainfall is inherently of short duration with high intensity, it is expected that the contribution of this type of rainfall will be stronger than in the past due to changing climate. While river flooding is mainly associated with extreme precipitation on the daily time scale, extreme sub-daily rainfall often is the primary driver of urban flooding or flooding in small and/or steep catchments. In addition, hourly rainfall is of the main interest of engineering practice for the design of infrastructure, which necessitates identifying the relationship between the rise of temperature and extreme precipitation (Alfieri *et al.*, 2017). Employing a time series of observed hourly rainfall from 1966-2013 over the Australian Continent, Guerreiro *et al.* (2018) explored the effect of a warming climate on extreme precipitation for both daily and sub-daily time scales. It has been found that the percentage of change corresponding with the daily time-scale is in line with the CC-scaling curve. However, the scaling factor for extreme sub-daily precipitation falls within 2CC and 3CC curves manifesting the heightened sensitivity of hourly compared to extreme daily precipitation.

Translation of rainfall to flood is, however, a complex issue. One reason is the difference in the definition of 'extreme' for precipitation and flood. For precipitation, the 95 or 99 percentiles are popular indices to calculate extremes. For floods, a much larger index is needed to satisfy the constraints corresponding to the life-cycle of structures such as dams, levees, or bridges (Westra *et al.*, 2014). Such structures with a life-cycle reaching several decades should be designed to stand the floods with at least a 100-year return period (Kuczera *et al.*, 2006). Moreover, the non-linear response of catchments toward forcing meteorological variables, which stem from their size, geo-morphological characteristics, soil type, land cover, level of urbanization, etc. is another reason that increases the ambiguity regarding the translation of extreme rainfall into flood events (Bennett, Leonard, Deng & Westra, 2018). The size of the catchment is an

important factor in determining the response of the catchment toward rainfall, as the time of concentration for each catchment is highly correlated with its size. The time of concentration is the duration in which the runoff generated from the most hydraulically remote part of the catchment departs the outlet (Blöschl, 2022b). The time of concentration is often considered as the critical duration of an extreme precipitation. This is the duration at which the largest flood is assumed to accrue for a given rainfall.(Chow, Maidment & Mays, 1988; Brutsaert *et al.*, 2005). The critical time of catchments with a main channel of less than 100 km is generally less than a day, making them sensitive to sub-daily rainfall duration (Sikorska, Viviroli & Seibert, 2018).

1.5.2   Spatial Scale

Spatio-temporal scale issues related to runoff generation is the subject of continuous debates (Blöschl *et al.*, 2019; Blöschl & Sivapalan, 1995). In their article about 23 unsolved problems in hydrology (UPH), Blöschl *et al.* (2019) raised multiple questions that highlight the importance of improving our understanding of hydrological processes at different time-space scales. Question 5 from the 'spatial variability and scaling' category focuses on hydrological variables' spatial heterogeneity and homogeneity:'What causes spatial heterogeneity and homogeneity in runoff, evaporation, subsurface water and material fluxes (carbon and other nutrients, sediments), and in their sensitivity to their controls (e.g. snow fall regime, aridity, reaction coefficients)?'. Question 6 reflects on the hydrologic laws at the catchment scales and their variation with scales: 'What are the hydrologic laws at the catchment scale and how do they change with scale?'. Further, question 10 from the category 'variability of extremes' highlights the link between catchment heterogeneity (land cover/land use) and extreme runoff: 'Why are runoff extremes in some catchments more sensitive to land-use/cover and geomorphic change than in others?'.

Scale issues are rooted in the aggregation procedure required to reduce the computational burden in hydrological modeling (Gupta, Rodríguez-Iturbe & Wood, 2012; Sivapalan & Blöschl, 2017; Blöschl, 2022a). However, scaling hydrological processes, from pore- to catchment-scale and beyond, causes hydrological variability (Dooge, 1986). Based on Dooge (1986), variability can be classified as organized simplicity, unorganized complexity, and organized complexity. The

first and second forms can be resolved by mechanistic and probabilistic approaches, respectively. The third form, into which most of hydrological problems fall, is difficult to resolve (Weinberg, 2001; Dooge, 1986; Freeze & Harlan, 1969). There is a debate on whether to consider variability as uncertainty, in the sense that randomness is an inherent part of hydrological processes or considering that it is a result of lack of knowledge, which by performing more accurate measurements can be resolved in a deterministic manner (Western, Blöschl & Grayson, 2001; Rajaram, 2016; Montanari & Koutsoyiannis, 2012; Rigon *et al.*, 2022). Those philosophies give rise to contrasting modeling approaches, i.e., deterministic or probabilistic, which highlights the challenges to connect scales. Since flood generation processes and their linkage across scales are key modeling concepts, they will be briefly discussed in the following.

At the pore scale (10^{-3}m), gravity and capillarity are dominant forces determining water movement (Jury & Horton, 2004). While these forces are well understood in theory, the interaction between water and soil can become complex. Clay grains, for example, tend to change size under wet and dry conditions, depending on their type and shape (Tessier, 1990; Tuller & Or, 2003; Alaoui, Rogger, Peth & Blöschl, 2018). Changing size further causes the emergence of cracks on the land surface. Cracks are pathways of preferential flow, which increase infiltration and decrease local surface runoff (Blöschl *et al.*, 2016). At the pore scale and plot scale (1m), water movement through soil cannot be fully described by the well established Richard's equation. The equation is driven by a combination of the Darcy-Buckingham law and continuity equation(Richards, 1931). Despite extensive use in hydrological modeling, Richard's equation has limits in the application for all types of soil (Blöschl, 2022a). Moreover, laboratory studies have shown that water tends to bypass the soil matrix through macro-pores (i.e., preferential flow); a runoff generation process that cannot be captured by this equation (Flury, Flühler, Jury & Leuenberger, 1994; Zehe, Elsenbeer, Lindenmaier, Schulz & Blöschl, 2007). The generation process at the hillslope scale (100m) is different from the profile scale (1m). The Hortonian infiltration excess is the dominant runoff generation process at the plot scale. At the hillslope level, however, the spatial connectivity redistributes subsurface lateral

flow and gives rise to saturation excess runoff generation process (Zehe, Blume & Blöschl, 2010; Silasari, Parajka, Ressl, Strauss & Blöschl, 2017).

Runoff generation processes at the catchment scale are a subject of ongoing debates. There are two conflicting approaches: The traditional approach is based on Freeze & Harlan (1969), which focuses on explaining individual hillslope responses and combining them to reach the catchment response. This is an upward approach that sees the catchment as an organized simplicity. The other approach sees the catchment as an unorganized complexity, in which simply aggregating micro-processes (here at hillslope or more minor scales) will not give an adequate explanation of macro-processes (here catchment scale) (Sivapalan & Blöschl, 2015). This is because such an approach fails to consider the evolutionary spatial patterns that develop surface and subsurface drainage systems across scales (Blöschl, 2022a). One can take a catchment as an ecosystem that is conditioned on wetness and available nutrients. The presence of life shapes patterns in the catchment drainage system in a complex manner and produces a non-linear response that cannot be resolved by mechanistic approaches (Savenije & Hrachowitz, 2017). Therefore, a more holistic approach is needed to treat the catchment response, for example, by comparing and contrasting multiple catchments and analyzing their responses (Sivapalan, 2003).

1.5.3   Representation of Scale in hydrological models

Recent progress in process conceptualization (Miyata *et al.*, 2019; Sidle, Gomi, Usuga & Jarihani, 2017; Zehe & Blöschl, 2004), data acquisition through remote sensing and reanalyses techniques can help improve representing spatio-temporal scale in hydrological modeling (Singh & Sinha, 2021; Jiang & Wang, 2019; Koci, Sidle, Jarihani & Cashman, 2020). Multiple datasets with fine spatial and temporal scales have been developed to address this problem. The examples are Shuttle Radar Topography Mission (SRTM), The Global Multi-resolution Terrain Elevation Data 2010 (GMTED2010- with spatial resolution of 7.5 arcseconds), and Moderate Resolution Imaging Spectroradiometer (MODIS), which provide topographic, land use, and snow cover data with with spatial scales of 250m to 1 km (Pham, Marshall, Johnson & Sharma, 2018; Danielson & Gesch, 2011). The European Centre for Medium-Range Weather Forecasts (ECMWF) proposed

multiple reanalysis datasets (ERA-Interim, ERA5, ERA-Land), which include time-series of climatic variables with fine spatial (10 km for ERA-Land) and temporal (hourly) resolutions (https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5). These datasets can incorporate the details of spatio-temporal variability in implementing of process-based distributed hydrological models (Tarek, Brissette & Arsenault, 2020b).

The effect of spatial discretization on models' response has been investigated for urban catchments. It is shown that the changes in resolution of topographic information provided by digital elevation models (DEMs), for instance, could reorient flow direction and flow accumulation and alter the surface and channel routing schemes (Cao *et al.*, 2020a; Krebs, Kokkonen, Valtanen, Setälä & Koivusalo, 2014). Furthermore, altering soil texture and land cover modifies imperviousness, Manning roughness coefficient, soil water content, etc., and reshapes the final response in terms of both water generation and routing schemes (Cao, Ni, Qi & Liu, 2020b). This might potentially determine the timing and magnitude of the flood. It is worth noting that there is no consensus over the impact of additional information provided by high spatial resolution on the estimation of high flows, as many studies show contradicting results (i.e., overestimation or underestimation of flood) by refining spatial resolution (Ichiba *et al.*, 2018; Warsta *et al.*, 2017).

Over-land flow connectivity is another important scale-related issue that affects catchment responses (Moore & Grayson, 1991; Jones, Swanson, Wemple & Snyder, 2000; Croke, Mockler, Fogarty & Takken, 2005). Several studies have addressed the impact of flow connectivity in water and nutrient transports across hillslope-riparian-river systems (Jencso, McGlynn, Gooseff, Bencala & Wondzell, 2010; Sedell, Reeves, Hauer, Stanford & Hawkins, 1990; Pringle, 2001). Different methods have been proposed to incorporate connectivity into hydrological models including contour line-based method (Moore & Grayson, 1991; Miyata *et al.*, 2019), topographic index methods (Detty & McGuire, 2010; Hallema, Moussa, Sun & McNulty, 2016), empirical approaches (López-Vicente *et al.*, 2017; Koci *et al.*, 2020), and graph theory approach (Halverson & Fleming, 2015; Cossart & Fressard, 2017). In process-based distributed hydrological models, surface connectivity is directly related to the resolution of topographic

and land cover data represented in the model. By changing the modeling scale, a degree of variability corresponding with modeling parameters is expected to emerge. In fact, the choice of spatio-temporal discretization can be seen as a form of aggregation over scale, which causes variability in simulating hydrological variables.

Distributed grid-based models offer the opportunity to represent spatial heterogeneity if refined data are available adequately. The major challenge, however, is the trade off between the computational cost of implementing such models and the additional information the model can provide by refining grids. Zhang & Montgomery (1994), for example, showed that by increasing grid resolution higher than a specific threshold, the model response has marginally been improved. Therefore, it is plausible to force the model with a coarse grid without risking the elimination of necessary information. As a result, lower computational cost accompanied by an adequately accurate response boosts the efficiency of the model. One way to avoid model complication and maintain accuracy at the same time is the representation of subgrid spatial variability through suitable parametrization (Clark *et al.*, 2015). As such, surrogate information of gridded data (land use, soil texture, topography) encapsulated in the parameters could be transferred across regions and scales (Fatichi *et al.*, 2016).

1.6   Regionalization

During the past two decades, parameter regionalization techniques have been used to facilitate transfer of information across regions (Götzinger & Bárdossy, 2007; Hundecha & Bárdossy, 2004) and scales (Kumar, Samaniego & Attinger, 2013; Samaniego *et al.*, 2010). In practice, most of the regionalization techniques available in the literature use parameterization at the catchment scale and overlook subgrid variability. This may cause problems in transferring information across scales (Troy, Wood & Sheffield, 2008). Furthermore, using discrete classes of catchment descriptors may create a discontinuous distribution of state variables and fluxes, even though the streamflow is accurately simulated (Hundecha & Bárdossy, 2004).

To address the above mentioned problems, Samaniego *et al.* (2010) and Kumar *et al.* (2013) proposed a method to account for the variability below the modeling scale through a technique called Multiscale Parameter Regionalization (MPR). In this technique, the finest available gridded data is used to develop a regionalized transfer function, and an appropriate operator is employed to upscale the transfer function to the modeling scale. The a priori relationships between catchment descriptors and model parameters are developed using pedotransfer functions, which are based on empirical evidence and expert knowledge. This technique is further embedded in meseo-scale hydrological model and applied to broad variety of problems (Rakovec *et al.*, 2016; Baroni, Zink, Kumar, Samaniego & Attinger, 2017; Thober *et al.*, 2019).

The fundamental feature of regionalization techniques, including MPR is to establish an a priori relationship between parameters and catchment descriptors (Hundecha & Bárdossy, 2004). However, the explicit relationship between model parameters and catchment descriptors is hard to detect (Merz & Blöschl, 2004; Parajka *et al.*, 2013) because the available data of catchment descriptors are represented as an average or percentage of those particular descriptors over the whole catchment (percentage of land use, average soil depth, etc.). As a result, catchment descriptors are not thorough representatives of catchment heterogeneity (Merz, Tarasova & Basso, 2020). In addition, the scale of the study can determine which groups of catchment descriptors are more important. For example, topography of the catchment could be a dominant descriptor at the catchment scale (Von Freyberg, Radny, Gall & Schirmer, 2014) while climate characteristics are more important at the regional scale (McGlynn & McDonnell, 2003).

The application of machine learning (ML) and deep learning (DL) in hydrology is rapidly increasing. In terms of type, most of the studies in the past were limited to the multi-layer perception (MLP) neural networks, and some used recurrent neural networks (RNN Abrahart *et al.*, 2012; Dawson & Wilby, 2001; Oyebode & Stretch, 2019). More recent studies have focused on the application of DL or, more precisely, Long Short-Term Memory (LSTM) network (Hochreiter & Schmidhuber, 1997), which is a special type of RNNs. LSTM has shown promise in hydrological modeling and regionalization (e.g. Kratzert, Klotz, Brenner, Schulz & Herrnegger, 2018; Kratzert *et al.*, 2019a; Gauch *et al.*, 2021a; Gauch, Mai & Lin,

2021b; Ali, Ebrahimi, Ashiq, Alasta & Azari, 2022; Duan & Ullrich, 2021; Li *et al.*, 2022; Mai, Craig, Tolson & Arsenault, 2022a; Arsenault, Martel, Brunet, Brissette & Mai, 2022). ML and DL can improve streamflow simulations and regionalization because they are typically trained on large datasets of highly variable catchments (Arsenault *et al.*, 2022). Despite their efficiency, ML and DL have limited use in the operational forecast community due to their black box nature. In parallel, the academic community continues to advance ML and DL applications, because contrary to conceptual models, these models have more degrees of freedom to numerically capture non-linear hydrological relationships (Nearing *et al.*, 2020; Kratzert *et al.*, 2019b). This helps transferability and scaling of hydrological relationships (Sidle, 2021).

A combination of hydrological models and ML/DL is a way forward. ML techniques are well suited to calculate the hydrological model parameters, for which a tangible physical relationship is not available. The advantages of ML regionalization methods have been demonstrated in the literature. Razavi & Coulibaly (2017) compared traditional methods with various forms of Artificial Neural Network (ANN) and ML methods in identifying regionalized parameters of a hydrological model for prediction in ungauged basins across Ontario. Their results showed that a certain combination of non-linear data-driven methods and catchment classification could potentially improve the performance of the model for PUB. Saadi, Oudin & Ribstein (2019) used Random Forest (RF) and two other traditional methods to regionalize the parameters for multiple urban and natural catchments across France and the US. Their results demonstrated that the RF method outperforms the traditional methods in both urban and natural catchments. Merz *et al.* (2020) developed Parameter Set Shuffling (PASS) method to develop a relationship between catchment descriptors and parameters using RF. PASS was successfully applied on multiple catchments in Germany to find consistent regionalized parameters over the country using a large-scale distributed hydrological model.

Regarding the application of ML and DL in regionalization, there is a gap about the representation of catchment heterogeneity and its importance for prediction in ungauged basins. The questions such as how well ML methods perform at different spatio-temporal resolution; to what extent the developed hydrological relationships between catchment descriptors and model parameters can

be transferred across scales; and to what extent using more spatially representative catchments helps improve the efficiency of the modeling have yet to be studied. Addressing these questions using ML/DL techniques will potentially help improve our understanding of scale issues in hydrological modeling.

1.7   Research Objectives and hypotheses

The main objective of this study is to assess the impact of spatial and temporal resolutions of distributed (and physically based) hydrological models on the simulation of flows, in particular summer-autumn floods. More specifically, three hypotheses will be examined:

1. The choice of level of spatio-temporal discretization alters model parameters, which leads to an uncertainty in streamflow and flood simulation. By increasing catchment area, the contribution of the choice of spatial scale and hydrological model in such uncertainty increases, and that of time-step decreases.

2. For small catchments ($< 500 km^2$), refining temporal resolution of simulation (from daily to subdaily) increases the relative change (from reference to future) of extreme summer-fall flow. Refining temporal resolution will not significantly affect projected extreme summer-fall flow for large catchments ($> 1000 km^2$).

3. There does exist relationships between model parameters and catchment descriptors that can be approximated by the random forest (RF) method. The underlying information related to catchment characteristics is transferable across scales through that approximation. Using fine-scale (in time and space) catchment descriptors improves the skill of the regionalization model.

CHAPTER 2

METHODOLOGY

This chapter provides a perspective on the methods and materials used in this research. The details of the experimental design corresponding with the research hypotheses can be found in chapters 3 to 5, and therefore, they are briefly addressed in this chapter.

2.1   Study Area

All catchments of this study are located in the Southern Quebec , Canada. The number of catchments varies, depending on the hypotheses. For the first and second hypotheses, we selected 6 and 4 catchments located in different hydrological regions of the province. Given that the main theme of the present study is to investigate the relationship between size and the spatio-temporal representation of the catchments in hydrological models, the main selection criterion is to include different catchment sizes. The sizes range from less than 200 to more than 3500 km^2. To verify the third hypothesis, we had to use the maximum possible number of catchments to introduce them to the RF model. In that case, we used 171 catchments, all located in southern Quebec. Figure 2.1 displays the catchments used in the experiment to verify the last hypothesis.

2.2   Data

We employ ERA5 precipitation and temperature time series (Hersbach & Dee, 2016, ECMWF ReAnalysis5;) to calibrate and validate the hydrological models. ERA5 is the fifth generation of ECMWF reanalyses with a spatial resolution of 31km and hourly temporal resolution. ERA5 has shown a good performance when compared with observed gridded datasets (Tarek *et al.*, 2020b). Observed streamflow series are obtained from the Direction de l'Expertise Hydrique (DEH) of the Ministère de l'Environnement et de la Lutte contre les changements climatiques (MELCCC) for the 2000-2017 time period, with 24 and 3-hour time steps.
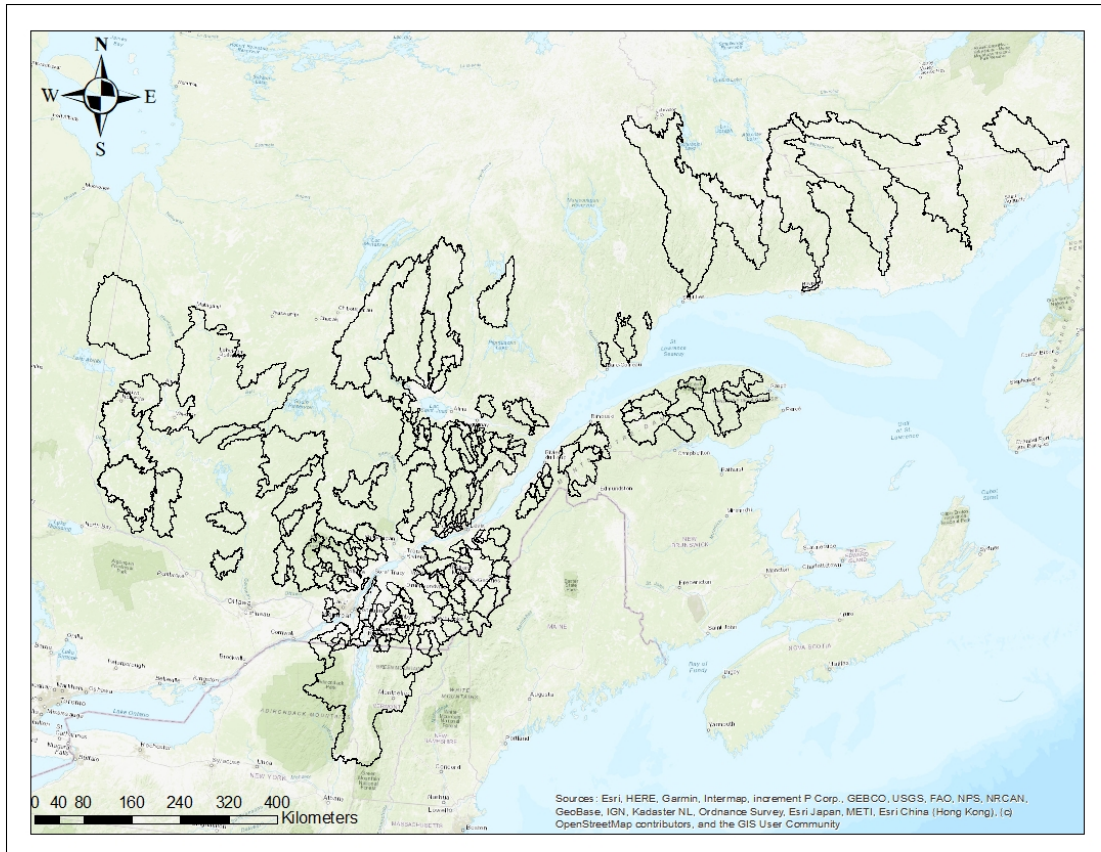
Figure 2.1   Location of the catchments used in this study, more specifically in the
case of hypothesis 3. All catchments are located in Southern Quebec, Canada

ClimEx large ensemble (Leduc *et al.*, 2019) dataset is used for projections. The dataset is a
50-member ensemble driven by a transient run of the second version of the Canadian Earth
System Model (CanESM2-LE at 200 km resolution; Swart *et al.*, 2019). The RCP 8.5 scenario
was used to run the model for the period covering 1951-2100, resulting in 7500 years of data,
with hourly time steps and a $11°$ spatial resolution.

2.3   Experimental setup

Figure 2.2 displays the experimental setup. The first and second columns deal with the first two
hypotheses regarding the uncertainty linked to the spatio-temporal variability in the simulation
and projection of the streamflow for the present and future periods. The last column shows

the random forest (RF) regionalization method associated with the third hypothesis. The first step calibrates the models with different spatio-temporal resolutions and creates an ensemble of simulations per catchment. This allows quantifying the uncertainties linked to catchment heterogeneity. The second step bias corrects a 50-member ClimEx dataset and forces the hydrological models to project streamflow in the 1991-2100 period at different spatio-temporal resolution. This follows by frequency analyses to estimate the change of summer-fall flood in the future and investigate the effects of spatio-temporal resolution, catchment size, and hydrological model on the variability of simulations. The third part of this thesis uses the Random forest method for the regionalization of a distributed hydrological model (Hydrotel) at different spatio-temporal resolutions. The random forest model creates a relationship between model parameters and catchment descriptors. This relationship is further applied to pseudo-ungauged catchments for model confirmation. Nested catchments are further identified and modeled to investigate parameter transferability of regionalization model across catchmnet, sub-catchment, and RHHUs.

2.3.1   Calibration

We used dynamically dimensioned search (DDS Tolson & Shoemaker, 2007) to calibrate the hydrological models (WaSiM and Hydrotel). DDS is a heuristic optimization algorithm that searches for a good global solution in the n-dimensional surface of the model parameters and the objective function. The algorithm relies on a user-specified budget of function evaluations. A probability is dynamically adjusted to limit the number of dimensions being perturbed according to the number of trials. This way, the number of trials can be determined based on available computational power. In each trial, the parameters are perturbed by scaling a random selection from a normal distribution with zero mean. The scale parameter is defined to allow the algorithm to escape from poor local optima. The default value of the scaling parameter is 0.2. Lower values can be used if a more refined search is required in the local region of the parameters (In all cases, we used the scaling factor of 0.2, as recommended by the documentation).
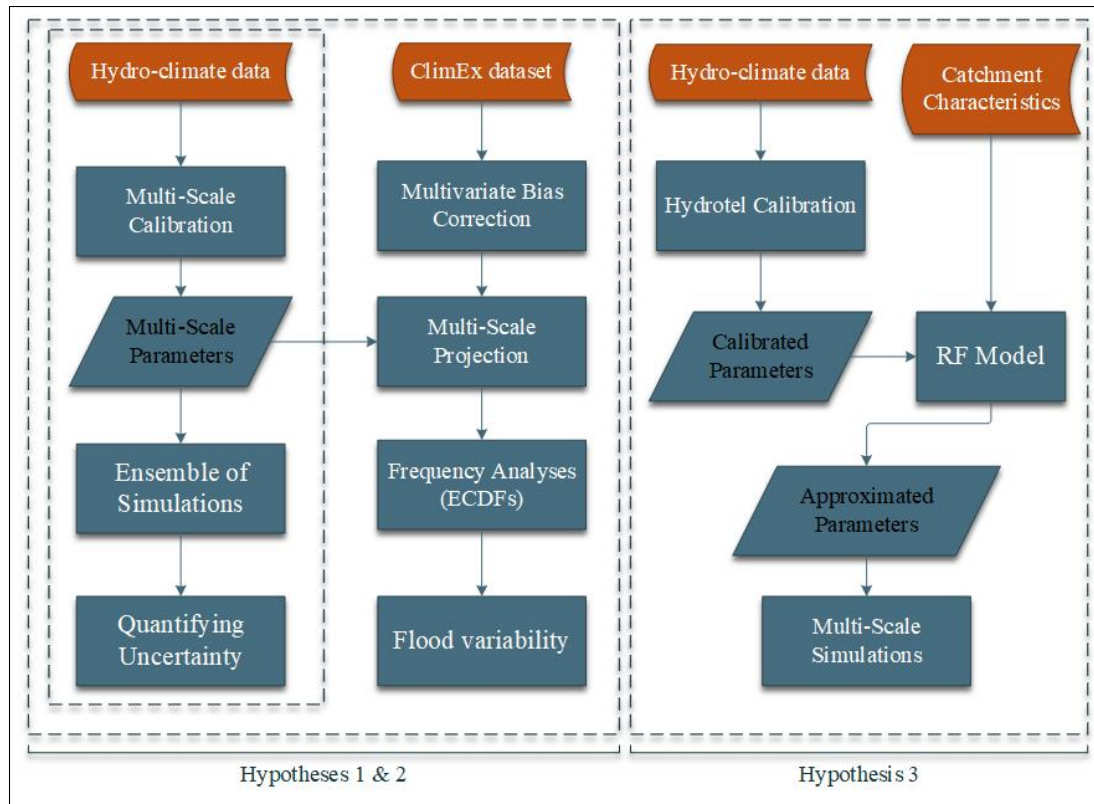
Figure 2.2   A brief schematic of the experimental design of this study

Multiple researches have used DDS to calibrate the parameters of hydrological models (e.g., Zhang, Srinivasan, Zhao & Liew, 2009; Arsenault, Poulin, Côté, Brissette *et al.*, 2014; Behrouz, Zhu, Matott & Rabideau, 2020; Darbandsari & Coulibaly, 2020). The algorithm has shown a good performance when compared with other optimization algorithms such as the shuffled complex evolution (Huot, Poulin, Audet & Alarie, 2019). The simplicity and performance of DDS, as well as the user-specified budget, make the algorithm an optimal choice for calibrating computationally-intensive process-based distributed models.

For calibration, a common practice is to use multiple objective functions or one function that is a composition of multiple measures, such as the mean squared error (MSE), Nash–Sutcliffe efficiency (NSE Nash & Sutcliffe, 1970), or Kling-Gupta efficiency (KGE Gupta *et al.*, 2009). NSE is, in fact, the MSE that is divided by the variance of the observation and subtracted from 1. Being dimensionless and scaled between (∞ , 1], it is a suitable measure to compare

model performances. For NSE, the modeling baseline is the observation mean. This is a problem for highly seasonal hydrological variables such as streamflow in snow-dominated basins (Schaefli & Gupta, 2007). There are also other known issues with the NSE, such as an underestimation of bias and variability. To resolve these issues KGE, which is the Euclidean distance in the Pareto front of the correlation coefficient (r), relative variability (α), and a measure for bias (β) (Equation 2.1) was proposed (Gupta *et al.*, 2009). This measure is a better alternative to measuring model performances in snow-dominated catchments (for all calibration runs, we used KGE as the objective function).

$$KGE = 1 - \sqrt{(r - 1)^2 + (\alpha - 1)^2 + (\beta - 1)^2} \tag{2.1}$$

2.3.2   Multivariate Bias Correction

The multivariate bias correction method is driven by N-dimensional probability density function transform (N-PDF; Pitie, Kokaram & Dahyot, 2005). N-PDF is an image processing technique for the transformation of information between two images. In this technique, the original multi-variable matrices are rotated by a random orthogonal function, and the rotated variables' marginal distribution are projected by quantile mapping. The extra rotation step linearly combines the variables for mapping. The traditional univariate techniques do not consider the dependency between multiple variables, which leads to inaccurate projections for impact studies (Rocheta, Evans & Sharma, 2014). The algorithm repeats the rotations and mapping steps until the distributions of the source and target datasets are the same. The MBCn algorithm (Cannon, 2018) adapts N-PDF for the bias correction of climate data following these steps: First, the datasets are rotated based on Equations 2.2 to 2.4.

$$\tilde{\mathbf{X}}_S^{[j]} = \mathbf{X}_S^{[j]}\mathbf{R}^{[j]} \tag{2.2}$$

$$\tilde{\mathbf{X}}_P^{[j]} = \mathbf{X}_P^{[j]} \mathbf{R}^{[j]} \tag{2.3}$$

$$\tilde{\mathbf{X}}_T^{[j]} = \mathbf{X}_T^{[j]} \mathbf{R}^{[j]} \tag{2.4}$$

where P, S, and T refer to climate model projection (future), climate model simulation (historical), and observation (historical). $\mathbf{X}$ is the matrix of climate variables, $\mathbf{R}$ is the orthogonal rotation matrix, and j is the iteration number. The next step is quantile delta mapping (QDM). In addition to mapping, the method treats the data points in the projections that lie outside the range of the historical simulated dataset. This is important, especially for decadal or multi-decadal projections with a strong climate change signal. The QDM transfer functions for temperature are presented in Equations 2.5 and 2.6.

$$\Delta(i) = x_P(i) - F_S^{-1}(F_P(x_P(i))) \tag{2.5}$$

$$\hat{x}_P(i) = F_T^{-1}(F_P(x_P(i))) + \Delta(i) \tag{2.6}$$

where $x(i)$ is i^{th} datapoint (here temperature), $\hat{x}(i)$ is corrected data point, and F and F^{-1} are empirical cumulative distributions and their inverse functions. The final step is to perform an inverse rotation according to Equations 2.7 to 2.9:

$$\mathbf{X}_S^{[j+1]} = \hat{\mathbf{X}}_S^{[j]} \mathbf{R}^{[j^{-1}]} \tag{2.7}$$

$$\mathbf{X}_P^{[j+1]} = \hat{\mathbf{X}}_P^{[j]} \mathbf{R}^{[j^{-1}]} \tag{2.8}$$

$$\mathbf{X}_T^{[j+1]} = \mathbf{X}_T^{[j]} \tag{2.9}$$

where $\hat{X}$ is the matrix after applying QDM. These steps are repeated until the simulation (X_S) matrix matches the observation (X_T) matrix. The MBCn algorithm takes an extra step to preserve the trend by reordering the elements of the mapped matrix according to the ordinal rank of the projection matrix (Cannon, Piani & Sippel, 2020).

2.4   Frequency Analyses

For the flood frequency analyses (chapter 3) from hydrological modelling using historical ERA5 data, we examined multiple distributions to fit the streamflow maxima. The Log-Pearson type III distribution (Equation 2.10, Griffis & Stedinger (2007)) appeared to be the best fit for this case study.

$$f_Q(q) = \frac{1}{|\beta|}(\frac{lnq - \tau}{\beta})^{\alpha-1} exp(-\frac{lnq - \tau}{\beta}) \tag{2.10}$$

where the Napierian logarithm of the random variable Q is X, which has a Pearson type III distribution, $\Gamma(\alpha)$ is the complete gamma function, $\alpha = \frac{4}{(\gamma_x)^2}$, $\beta = \frac{(\sigma_x \gamma_x)}{2}$, and $\tau = \mu_x - 2(\sigma_x/\gamma_x)$, and μ_x, σ_x, and β_x are the mean, standard deviation and skewness coefficient, respectively.

To estimate streamflow values for the summer-fall floods from hydrological modelling using the ClimEx dataset (chapter 4), we pooled the large-ensemble streamflow annual maxima. The underlying assumption for pooling is that the period under study is stationary (Martel, 2019). We examined the trend at each data point using the Mann–Kendall test at a 95% confidence level. For all the catchments, no significant trend has been observed for the future (2081-2100) and historical (1991-2010) periods. In the next step, we sorted the long time series to create an empirical cumulative distribution function and calculate flood with different return periods. The

advantage of this method is to eliminate the uncertainty related to the choice of a parametric

distribution (Meresa & Romanowicz, 2017).

CHAPTER 3

SPATIO-TEMPORAL DISCRETIZATION UNCERTAINTY OF DISTRIBUTED HYDROLOGICAL MODELS

Siavash P. Markhali[1] , Annie Poulin[1] , Marie-Amélie Boucher[2]

[1] Department of Construction Engineering, École de technologie supérieure, 1100 Notre-Dame West, Montréal, Québec, Canada H3C 1K3
[2] Civil and Building Engineering Department, Université de Sherbrooke, 2500 Bd de l'Université, Sherbrooke, Québec,Canada J3X 1S1

Abstract

Quantifying the uncertainty linked to the degree to which the spatio-temporal variability of the catchment descriptors (CDs), and consequently calibration parameters (CPs), represented in the distributed hydrological models and its impacts on the simulation of flooding events is the main objective of this paper. Here, we introduce a methodology based on ensemble approach principles to characterize the uncertainties of spatio-temporal variations. We use two distributed hydrological models (WaSiM and Hydrotel) and six catchments with different sizes and characteristics, located in southern Quebec, to address this objective. We calibrate the models across four spatial (100, 250, 500, 1000 m) and two temporal (3 hours and 24 hours) resolutions. Afterwards, all combinations of CDs-CPs pairs are fed to the hydrological models to create an ensemble of simulations for characterizing the uncertainty related to the spatial resolution of the modeling, for each catchment. The catchments are further grouped into large ($> 1000 \text{ km}^2$), medium (between 500 and 1000 km^2) and small ($< 500 \text{ km}^2$) to examine multiple hypotheses. The ensemble approach shows a significant degree of uncertainty (over 100% error for estimation of extreme streamflow) linked to the spatial discretization of the modeling. Regarding the role of catchment descriptors, results show that first, there is no meaningful link between the uncertainty of the spatial discretization and catchment size, as spatio-temporal discretization uncertainty can be seen across different catchment sizes. Second,

the temporal scale plays only a minor role in determining the uncertainty related to spatial discretization. Third, the more physically representative a model is, the more sensitive it is to changes in spatial resolution. Finally, the uncertainty related to model parameters is larger than that of catchment descriptors for most of the catchments. Yet, there are exceptions for which a change in spatio-temporal resolution can alter the distribution of state and flux variables, change the hydrologic response of the catchments, and cause large uncertainties.

3.1   Introduction

Understanding the spatio-temporal scale of the representation of hydrological processes, and confronting the issue of scale mismatch within inter-connected hydrological units are two major challenges in hydrological modeling (Blöschl & Sivapalan, 1995; Blöschl *et al.*, 2019; Fatichi *et al.*, 2016; Beven, 2011). To better understand the complexity (heterogeneity) in hydrological systems, which is present under continuous internal change (e.g., land use change) and boundary conditions (e.g., changing climate), distributed hydrological models have been used across different spatio-temporal scales (Addor *et al.*, 2014; Blöschl, Reszler & Komma, 2008; Famiglietti & Wood, 1995; Kumar, Samaniego & Attinger, 2010; Kumar *et al.*, 2013; Merz & Blöschl, 2004; Wanders & Wada, 2015; Rakovec *et al.*, 2016; Martel, Brissette & Poulin, 2020a; Thober *et al.*, 2019). However, the models themselves suffer from inadequate simulation of hydrological processes due to a lack of scale-relevant theories in catchment hydrology (Samaniego *et al.*, 2017; Blöschl & Sivapalan, 1995; Dooge, 1986; Peters-Lidard *et al.*, 2017). In fact, changes in the spatio-temporal discretization of the physiographical characteristics of a catchment can alter the dynamic interactions between state variables and fluxes, resulting in different model responses (e.g., Cao *et al.*, 2020a; Krebs *et al.*, 2014). Therefore, part of the modeling uncertainty is due to the extent to which the physiographic characteristics of the catchment are described, more or less finely, by the model. Such uncertainty is normally ignored in practice, and is the focus of the present research. Specifically, we aim to quantify the relative roles of the spatial resolution of the physiographic characteristics, as well as that of the model's parameters obtained by calibrating the model using different spatio-temporal representations of

catchments. To this end, two different distributed hydrological models will be used, as well as six catchments, all grouped into an ensemble-based approach (Krzysztofowicz, 2001), involving 16 simulations per model and per catchment.

Unlike to lumped models, which treat the whole catchment as a unique homogeneous area, distributed models incorporate the spatial heterogeneity of the catchments. Depending on the level of discretization, distributed models can be classified into two broad categories: semi-distributed and fully distributed (Clark *et al.*, 2015, 2017). In semi-distributed models, of which SWAT (Arnold *et al.*, 1998) and VIC (Liang *et al.*, 1994) are two well-known examples, the level of spatial discretization is limited to defining the number of Hydrological Response Units (HRU). On the other hand, models such as WaSiM (Schulla & Jasper, 2007), MIKE-SHE (Refsgaard, 1995) and HYDRUS-3D (Šimunek *et al.*, 2008) are considered as fully distributed, as instead, they discretize the catchment using grids, and the computation of the fluxes and state variables is performed for each grid cell. Distributed models can also be viewed based on a physical or conceptual representation of the processes. Physically based models attempt to solve the conservation of mass, energy and momentum equations to represent hydrological processes at micro-scale control volumes (Hrachowitz & Clark, 2017; Fatichi *et al.*, 2016). MIKE-SHE (Refsgaard, 1995) and HYDRUS-3D (Šimunek *et al.*, 2008) are typical examples. Conceptual models represent processes more simply, through macro-scale conceptualization(Clark *et al.*, 2017; Devia, Ganasri & Dwarakish, 2015). The distributed version of the HBV model (Bergström *et al.*, 1995), mHM (Samaniego *et al.*, 2010) as well as CEQUEAU (St-Hilaire *et al.*, 2015a) can be placed in this category.

In flood forecasting, analyses of hydrological processes, or in climate change impact assessment studies, the underlying assumption for implementing a specific model over different spatio-temporal resolutions, is usually that the parameters are scale-invariant, ensuring the production of similar states and fluxes regardless of the spatio-temporal resolution (Samaniego *et al.*, 2017). However, such assumption is questionable in the absence of scale-relevant theories for natural catchments, as the heterogeneity of the system dominates the consistency needed across different catchments to develop a general theory (Hrachowitz *et al.*, 2013; Nearing *et al.*, 2020). In fact,

different hydrological processes that take place under different spatio-temporal scales at different catchments highlight the "uniqueness of the place" (Beven, 2000), as opposed to the generality of hydrological response. The problem is that the lack of such scale-relevant theories directly affects modeling practices. Model parameters, for example, typically represent hydrological processes that are either complex, or take place on a very small scale, or that are not yet well understood (Pokhrel & Gupta, 2010; Barrios & Francés, 2012; Brynjarsdottir & OHagan, 2014). In practice, for most cases, model parameters lack physical reality, as very often, there are no tangible links between catchment attributes and parameters(Beven, 1995). Furthermore, the dearth of knowledge regarding upscaling theories and their application in hydrological modeling exacerbates the problem (Neuman, 1990; Kitanidis & VoMvoris, 1983). Therefore, the parameters cannot be considered scale-invariant and the conditions of flux-matching across diverse spatio-temporal scales cannot be satisfied with current knowledge (Wood, Sivapalan, Beven & Band, 1988).

The randomness of hydrological processes, attributable to a lack of knowledge related to the complexity of the system, can be addressed by replacing the deterministic results of modeling with an ensemble of simulations using probabilistic or deterministic approaches (Dooge, 1986; Beven, 2006; Nearing, Gupta & Crow, 2013; Nearing & Gupta, 2015; Nearing *et al.*, 2020). We suggest that the principles of ensemble simulations can also be useful in addressing the uncertainty linked to the spatio-temporal variability of the physical descriptors of a catchment. As such, an ensemble of simulations derived from variations of CDs-CPs resolutions can be constructed for each catchment to quantify the uncertainties corresponding to the spatio-temporal resolution of the modeling. While multiple studies focus on accounting for and quantifying different sources of uncertainties in hydrological modeling, some include input data uncertainty, structural uncertainty, parametric uncertainties, or a combination of the preceding (e.g., Dixon & Earls, 2012; Faramarzi *et al.*, 2013; Zhao *et al.*, 2018; Poulin *et al.*, 2011; Joseph, Ghosh, Pathak & Sahai, 2018; Thiboult *et al.*, 2016; Euser *et al.*, 2013; Craig *et al.*, 2020; Tarek, Brissette & Arsenault, 2020a; Refsgaard, Van der Sluijs, Brown & Van der Keur, 2006; Butts *et al.*, 2004), less attention has been directed towards the uncertainty related to spatio-temporal

variability and how it impacts modeling. This may be attributable to a belief that such uncertainty has but trivial impacts on the modeling. However, among the limited research works that have been conducted in this context, Tegegne, Kim, Seo & Kim (2019) demonstrated that changing the sub-basin spatial scale in the SWAT model has a small impact on the entire flow simulations, but that a substantial sensitivity could be observed when reproducing more extreme flow quantiles. Their study, however, was limited to varying the number of HRUs, as opposed to changing the spatio-temporal discretization of the model's parameters. Moreover, no mechanisms were considered to account for the uncertainties related to spatio-temporal variability of the physical descriptors of a catchment.

Varying the spatial resolution used to represent land use in the model might also lead to a range of simulations, and therefore help to quantify the corresponding uncertainty. Distributed models have widely been used to account for land use change across the globe (e.g., Singh *et al.*, 2015; Li *et al.*, 2019; Yang, Long & Bai, 2019a; Tavangar, Moradi, Massah Bavani & Gholamalifard, 2019). In a series of papers (Breuer *et al.*, 2009; Huisman *et al.*, 2009; Viney *et al.*, 2009; Bormann, Breuer, Gräff, Huisman & Croke, 2009) under the project on 'Assessing the impact of land use change on hydrology by ensemble modeling (LUCHEM)', an ensemble of 10 hydrological models were used, with a range of structural complexity. More recently, Chen *et al.* (2019b) investigated parameter uncertainty stemming from land use change across different time-scales. They used two distributed models and three land use scenarios to simulate streamflow on a catchment located in China. Their results suggest that land use change does not have substantial effects on runoff simulations, but a large range of uncertainty can be observed for extreme streamflow values. It is worth noting that these research works focus on land use change scenarios, while the impact of change of spatio-temporal resolution on the modeling and the uncertainties are yet to be investigated.

The impact of spatial discretization on flood events has been investigated with a focus on urban catchments (e.g., Cao *et al.*, 2020a; Krebs *et al.*, 2014; Zhou *et al.*, 2017). It was found that changes in resolution of the topographic information provided by digital elevation models (DEM), for instance, could reorient the flow direction and flow accumulation, and alter surface

and channel routing (Cao *et al.*, 2020b). Furthermore, altering soil textures modifies the imperviousness, the Manning coefficient, the soil water content, etc., in addition to reshaping the final response in terms of both runoff generation and routing processes (Cao *et al.*, 2020a). Given the high degree of imperviousness and the complexity of surfaces in urban catchments, changes in spatial resolution could affect the results of flood simulations, which may leave such catchments more vulnerable to flooding events (Zhou *et al.*, 2017). Furthermore, changes in model response due to the degree to which the spatial heterogeneity of the catchment is represented might potentially affect the simulation in terms of peak timing and magnitude (Ichiba *et al.*, 2018). However, there is still no consensus on the impacts of refining the spatial resolution, as many studies show contradictory results, i.e., overestimation or underestimation of extreme flows (Warsta *et al.*, 2017).

While it has been shown that the choice of a particular level of spatio-temporal discretization on streamflow simulation of urban catchments can affect the simulation of peak streamflow, less focus has been given to similar issues for natural catchments. Still, the choice of a particular spatio-temporal resolution can similarly impact the modelling of natural catchments, since any variation of the land surface and slopes leads to a change in the time of concentration (Grimaldi, Petroselli, Tauro & Porfiri, 2012). This translates in variation of time and magnitude of flooding events. Such variations could also be important for climate change impact assessment studies, particularly for snow dominated catchments. Due to a changing climate, flooding patterns are shifting from spring to summer-fall (Kharin *et al.*, 2018). While floods due to the spring freshet are characteriezed by temporal scales covering several days, summer-fall flood events are usually characterized by much shorter temporal scales (sub-daily), to which small catchments are more sensitive (Yin *et al.*, 2018; Donat *et al.*, 2016). Consequently, depending on the catchment size, the choice of a specific spatio-temporal discretization is important for accurate flood simulations . Therefore, the choice of spatio-temporal discretizations in natural catchments needs to be further investigated. The respective roles of catchment area and characteristics, the time step of the simulation, as well as the model structure and parameters, are potentially important

determinants of a hydrological model's response, and this paper aims at investigating their roles. More specifically, we propose to test the following hypotheses:

1. Larger catchments are susceptible to larger uncertainties in the simulation of streamflow, when varying the spatial resolution of their physiographic characteristics.

2. Finer time steps introduce a higher degree of variability in the simulation, leading to increased uncertainty in streamflow simulation.

3. The more finely distributed and physically realistic a model is, the more sensitive to changes in spatial resolution it is.

4. The uncertainty related to model parameters is larger than that of catchments descriptors (DEM resolution, land use, soil texture).

These hypotheses will be examined through multiple experiments performed using two distributed models and six catchments of various sizes.The experiments will result in an ensemble of simulations to be investigated per catchment and per model. The structure of the paper is as follows. Section 3.2 provides details about the study area and the characteristics of the selected catchments, a brief description of the models used for simulations and the details of the experimental design. Results are presented in section 3.3 and discussed in section 3.4, taking one specific catchment as a representative example. Finally, concluding remarks and perspectives for future work are presented in section 3.5.

3.2 Method and Data

3.2.1 Study Area

Six catchments ranging from 100 km^2 to more than 2500 km^2 located in Quebec, Canada, are selected for this study (see Figure 3.1). The selection procedure is based on the following criteria: First, a broad range of catchment sizes should be covered to analyze the sensitivity of hydrological responses to the catchment size. Second, catchments should not belong to the same hydrological region, but rather, should be distributed across the territory (here the province of Quebec). Third, at least 10 years of streamflow data for 24- and 3- hour time steps need to

be available to fulfill the calibration and validation procedures. Table 3.1 describes the main characteristics of the catchments used in this study, which are identified in Figure 3.1. The catchments are sorted in descending order based on their area.
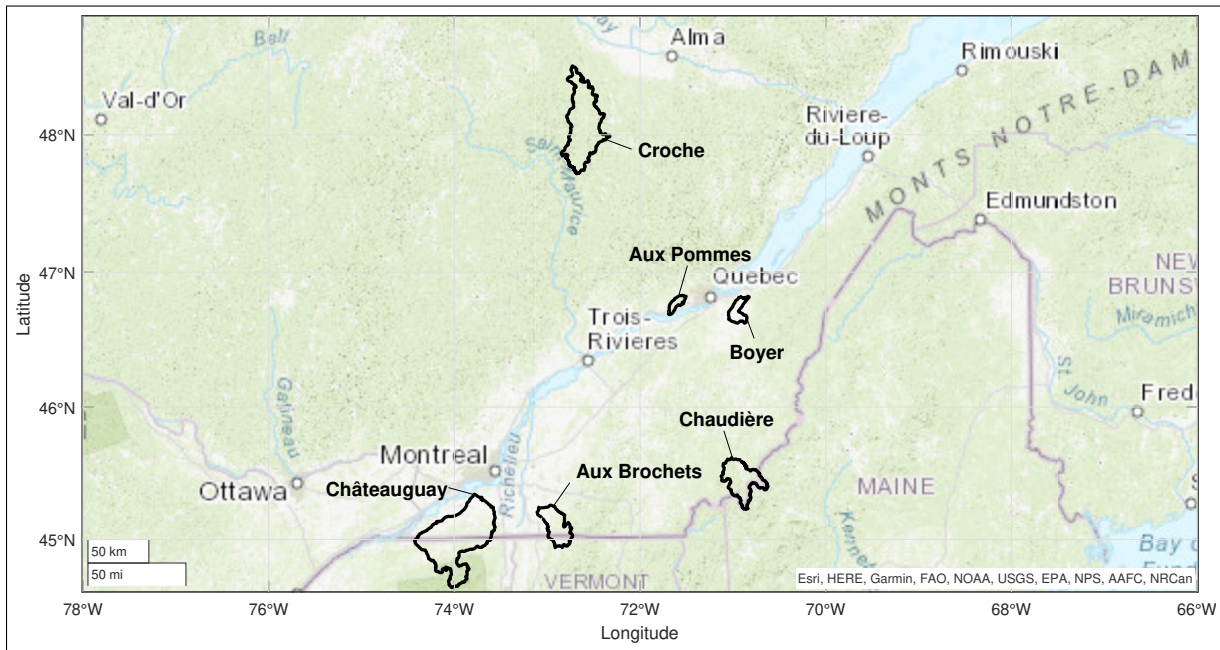
Figure 3.1 Location of the catchments for this study, over the southern part of Quebec

Table 3.1 General information and characteristics of the catchments

Station Number	Station Name	Area (km^2)	Mean annual precipitation (mm/yr)	Mean annual streamflow (m^3/s)	Mean annual temperature (°C)
030905	Châteauguay	2492	1192.77	38.45	6.65
050135	Croche	1563	1139.36	30.70	2.74
023427	Chaudière	781	1208.65	16.47	3.72
030424	Aux Brochets	584	1329.34	10.52	6.23
023004	Boyer	191	1396.76	4.45	4.15
050812	Aux Pommes	97.8	1322.69	2.37	5.12

3.2.2 Hydrometeorological data

The present study employs meteorological data (i.e. precipitation and temperature) extracted from ERA5 (ECMWF ReAnalysis5) gridded dataset to force the hydrological models for the historical time-period. Gridded reanalyses datasets are considered as an alternative to observed historical meteorological data. Using such datasets allow to solve major flaws of observational datasets, including missing data (particularly for higher resolutions), measurement errors, uneven distributions, etc. (Tarek *et al.*, 2020b). The European Centre for Medium-Range Weather Forecasts (ECMWF) proposed multiple reanalysis datasets (ERA-Inerim, ERA5, ERA-Land), which are widely used by hydro-climate modelers (Belmonte Rivas & Stoffelen, 2019; Wang, Graham, Wang, Gerland & Granskog, 2019). ERA5 is the fifth generation of ECMWF reanalyses of global climate products. The spatial resolution of ERA5 is 31km and the temporal resolution is hourly. Currently, the dataset covers the period from 1979 to today, and is expected to be updated to 1950 in the near future.

Observed streamflow series are obtained from the Direction de l'Expertise Hydrique (DEH) of the Ministère de l'Environnement et de la Lutte contre les changements climatiques (MELCCC) for the 2000-2017 time period, with daily and 3-hour time steps.

3.2.3 Hydrological models

3.2.3.1 WaSiM

The Water balance Simulation Model (WaSiM; Schulla & Jasper, 2007) is a process-based model that operates on a raster (grid) system. Its submodels run each grid cell of a catchment for each time step, providing the opportunity to use parallel computation algorithms based on the OpenMP standard. The model represents hydrological processes through its submodel structure, in which several options for interpolation, evapotranspiration, snow accumulation and melt, interception, glacier model, silting-up, unsaturated zone including heat transfer, saturated zone, surface discharge routing, and discharge routing including lakes and reservoirs are available.

The distinguishable feature of WaSiM is its provision of options to calculate infiltration and to represent water in the soil layers, with the calculation being more detailed than for most surface hydrological models. Two methods can be used namely, the modified conceptual Topmodel approach, and Richard's Equations approach (or unsaturated zone model). Since the second approach is more physically-based, we selected this version for simulations. The 1-D Richards equation, which represents fluxes in the unsaturated zone, is represented by Equation 3.1 (Schulla & Jasper, 2007):

$$\frac{\partial \Theta}{\partial t} = \frac{\partial q}{\partial z} = \frac{\partial}{\partial z}(-k(\Theta)\frac{\partial \Psi(\Theta)}{\partial z}) \tag{3.1}$$

where $\Theta(m^3/m^3)$ is the water content, $t(seconds)$ is time, $k(m/s)$ is the hydraulic conductivity, $\Psi(m)$ is the hydraulic head, $q(m/s)$ is the flux, and $z(m)$ is the depth of the soil column. WaSiM solves Equation 3.1 for multiple soil layers (the default is 30 layers for each type) of a grid cell using the finite difference method.

The unsaturated zone model controls multiple hydrologic variables such as infiltration, exfiltration, interflow, baseflow, real evapotranspiration, groundwater recharge, etc. Given the physical approach adopted to represent the flux of water in soil, WaSiM leans towards physically-based models. However, considering the simplified 1-D version of the continuity equation (instead of 3-D), and the existence of other empirical elements in the submodels (e.g., potential evapotranspiration) hinders the classification of the model among full physically-based distributed models. Table 3.2 specifies the choices that were made for each submodel of WaSim and for Hydrotel, which are described in the next sub section.

We calibrated 12 parameters: 7 for the infiltration model based on Richard's equation, 2 for evapotranspiration, 1 for the snow accumulation and melt model, and 2 for the spatial interpolation. Those parameters were selected after a consultation with the team who develops and maintains WaSim as well as from the WaSiM documentation. We set the remaining parameters to their default values, as per recommendation of the WaSiM documentation (Schulla & Jasper, 2007).

3.2.3.2   Hydrotel

Hydrotel is an HRU-based distributed model that is widely used operationally for flood forecasting by the DEH (e.g., Martel *et al.*, 2020a; Turcotte, Morse & Pelchat, 2020; Lucas-Picher *et al.*, 2020). The model adopts a mixture of physical, conceptual and empirical relationships to represent hydrological processes. Like WaSiM, it provides multiple options for calculating the hydrological processes of a catchment. The main particularity of Hydrotel is its compatibility with GIS and remotely sensed data (Fortin *et al.*, 2001a). Therefore, the model is capable of representing the spatial variability and the topography of catchments through a digital elevation model (DEM), soil texture maps and land use data through its components.

The model uses BV3C (Bilan Vertical 3 Couche) for soil modeling, which is specifically developed for Hydrotel. In this approach, the soil column is divided into three layers: The first layer is a surface layer that controls infiltration and is affected by surface evaporation; the second layer is associated with interflow, and the third one controls the baseflow. For the whole soil column, a moisture accounting equation is designed to represent macroprocesses of fluxes (Fortin *et al.*, 2001a). As a result, from a model classification perspective, the model leans towards the group of conceptual, distributed models, even though Hydrotel comprises certain physically-based elements related to surface and channel routing. Table 3.2 shows the submodels of Hydrotel used in this study for simulations.

Hydrotel has 28 parameters, of which eleven were calibrated. This includes 3 parameters for the infiltration and percolation submodel (BV3C), 6 for the snow accumulation and melt submodel, 1 for potential evapotranspiration, and one for the spatial interpolation (Thiessen polygons). The remaining parameters were fixed based on previous studies from our group, from the current practices at the government of Quebec (where Hydrotel is used operationally) and from the available documentation (see Huot *et al.* (2019) for more details about the parameters).

Table 3.2   The submodels employed to represent the hydrological processes in Hydrotel and WaSiM

Submodels	Hydrotel	Wasim
Interpolation	Thiessen polygons	Thiessen polygons
Snow melt/accumulation	Degree-Day Method	Degree-Day Method
Potential evapotranspiration	Hydro-Quebec Fortin *et al.* (2001b)	Hamon Hamon (1961)
Real evapotranspiration	after BV3C	after Richards' Eq.
Soil model	BV3C	Richards' Eq.
Channel routing	Kinematic Wave Eq.	Kinematic Wave Eq.

It should be noted that we developed two types of configurations for the simulations with Hydrotel, in order to allow the comparisons between a grid-based model (i.e., WaSiM) and an HRU based model (i.e., Hydrotel). In the first configuration (referred to as Hydrotel1 hereafter), we keep the number of HRUs constant, while the spatial resolution varies. Since the number of HRUs are kept constant during this experiment, the model might show low sensitivity to changes in spatial discretization. To rectify this, we introduced the Hydrotel2 experiment. In this configuration, we adjust the number of HRUs to match the change in resolution. We manually set the number of HRUs equal to the number of subbasins, which are automatically created for WaSiM based on the spatial resolution of CDs. This configuration is referred to as Hydrotel2 hereafter.

3.2.4   Experimental plan

Figure 3.2 delineates the different steps of our methodology and the experiments designed to answer the question posed in the introduction. The first column of the figure shows the "Data Domain", comprised of forcings (precipitation-temperature), calibration data (observed streamflow), and gridded Catchment Descriptors (CDs- e.g., DEM, land use, soil texture). For CDs, the highest available resolution is 100 m and we used resampling and interpolation methods to upscale the grids to 250 m, 500 m, and 1000 m resolutions. The second column, which is

referred as "time domain" shows the time step of forcing and calibration data. For this project, the subdaily time step is equal to 3 hours.

Regarding the third column titled "Calibration", as per usual, we split the time-series into calibration and validation periods. The duration of both periods are equal unless there exists a large part of missing data in between them that could reduce the accuracy of the calibration. In such a case, we remove that specific year from the period of calibration or validation. It is worth mentioning that the time-series of data related to winter streamflow in 3-hour time step is not available, and as a result, we removed this part of the year from the analyses.
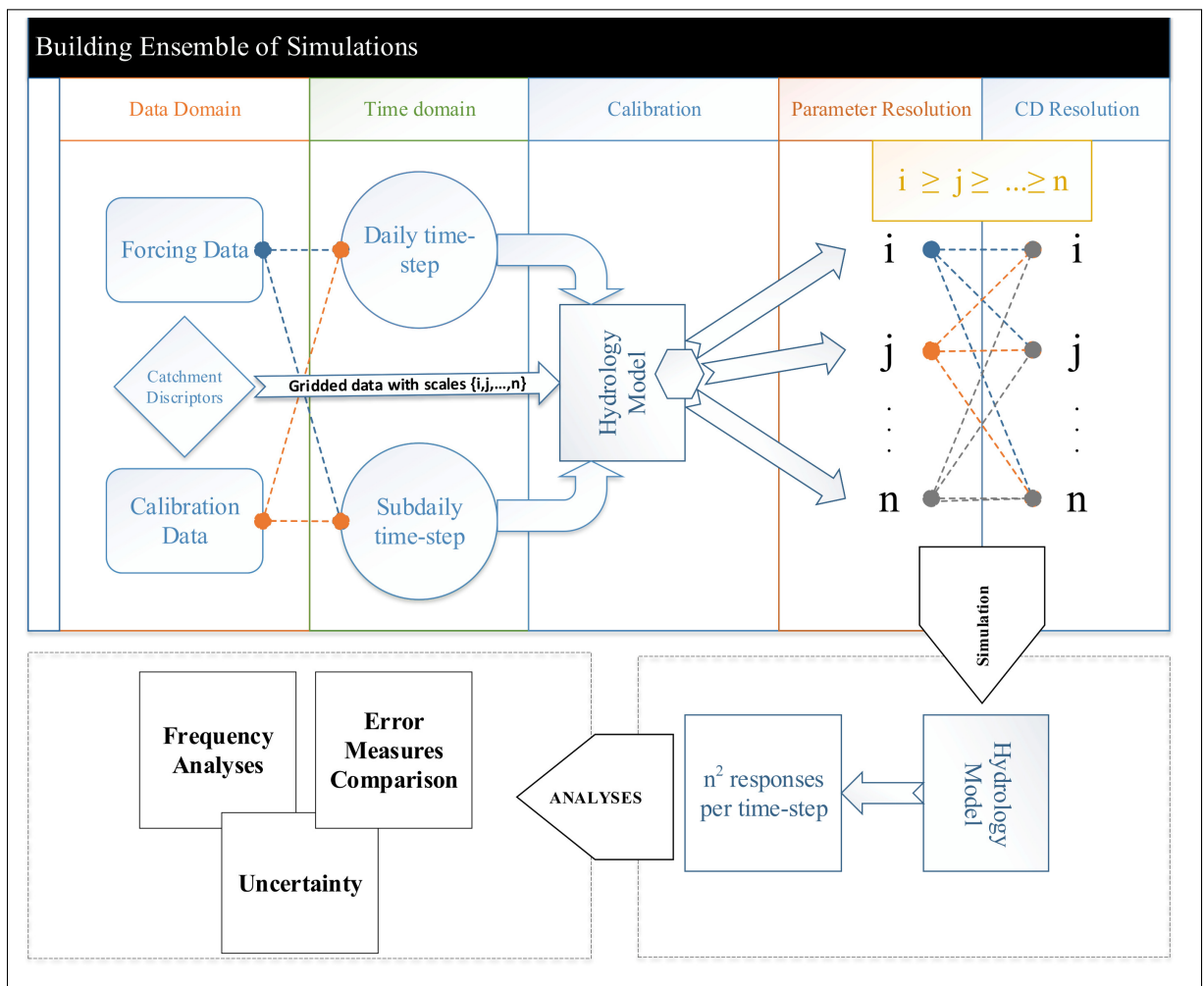
Figure 3.2   Schematic explanation for building ensemble of simulations and analyses

We used the Dynamically Dimensioned Search (DDS; Tolson & Shoemaker, 2007) algorithm to calibrate the hydrological models. The DDS global approach to scale the search for parameters based on the user-specified budget is an advantage compared to many other algorithms. In addition, its global approach to perturb parameters at the beginning of the search, which then narrows down the search space when the calibration procedure progresses, was shown effective in the literature (e.g., Huot *et al.*, 2019). Since most of the work for this study was carried out in Matlab, the main function of DDS was also executed in Matlab. Multiple scripts were developed to facilitate the communication between the models and the main DDS algorithm. A 0.2 perturbation factor has been used for the optimization. Furthermore, the Kling-Gupta Efficiency (KGE; Gupta *et al.*, 2009) is adopted as the objective function for optimizations. Compared to the Nash–Sutcliffe efficiency, KGE is more suitable to compare modeling skills across various catchments and time-steps with multiple variability. Moreover, KGE provides a more realistic view regarding the modeling efficiency for snow-dominated catchments (Gupta *et al.*, 2009). The KGE is computed using Equation 3.2.

$$KGE = \sqrt{(r-1)^2 + (\frac{\sigma_{sim}}{\sigma_{obs}} - 1)^2 + (\frac{\mu_{sim}}{\mu_{obs}} - 1)^2} \tag{3.2}$$

where r is the linear correlation between observations and simulations, σ_{sim} is the standard deviation in observations, σ_{obs} is the standard deviation in simulations, μ_{sim} is the simulation mean, and μ_{obs} is the observation mean.

When the distributed models are fed and calibrated against streamflow at the outlet of the catchment, several calibration parameter sets are obtained according to the spatio-temporal discretization of the input data (forth and fifth columns titled "Parameter Resolution" and "CD Resolution"). In the next step, all combinations of CPs-CDs are used to force both hydrological models for simulations. With n = 4 different resolutions for each calibration, an ensemble of n^2 = 16 simulations is obtained for each model (i.e. WaSiM, Hydrotel1, and Hydrotel2) and catchment.

To explore the uncertainty due to the spatial discretization, we first separate the catchments based on their surface areas to investigate the possible relations between discretization uncertainty and catchment size. Catchments are separated into three categories: larger than 1000 m (hereafter "large"), between 500 m and 1000 m (hereafter "medium"), and less than 500 m (hereafter "small"). As shown in Table 3.1, each category comprises two catchments. To find the variabilities corresponding with the change of spatial scales, we calculate the streamflow quantiles (i.e. $Q10$, $Q50$, $Q90$, and $Q95$) per simulation for each catchment in 24- and 3-hour time-scales separately (16 quantiles per model, catchment and time-step). Then we derive the variance of flow quantiles divided by the corresponding observed flow quantile and standardized them between zero and one. This enables us to compare the results among models, catchments and time-steps. Second, we compare the efficiency of simulations in calibration and validation across different spatio-temporal resolutions and explore the sensitivity of the efficiency of simulations to the changes in the CPs' and CDs' resolution. Third, we apply extreme value theory Coles, Bawa, Trenner & Dorazio (2001) to simulate flood events with different return periods by fitting the Log-Pearson distribution to the annual flow maximas. We calculate summer-fall floods with 5, 10, 20, and 50 years return periods for each simulation and calculate the relative error in flood simulations according to Equation 3.3:

$$e_{T,ij} = \frac{QT_{ij} - QT_{obs}}{QT_{obs}} \tag{3.3}$$

where e is the relative error of simulations, i is the CP resolution, j is the CD resolution, QT is the magnitude of a flood event with return period T, and *obs* represents the observation.

Given the 16 possible combinations of simulations, a range of relative error is obtained from Equation 3.3 for a specific return period. This range can further be separated into uncertainties corresponding to CPs and CDs according to Equations 3.4 and 3.5:

$$MDE_{T,i}^{CD} = |\max(e_{T,ii}, e_{T,ij}, ..., e_{T,in}) - \min(e_{T,ii}, e_{T,ij}, ..., e_{T,in})| \tag{3.4}$$

$$MDE_{T,j}^{CP} = |\max(e_{T,ij}, e_{T,jj}, ..., e_{T,nj}) - \min(e_{T,ij}, e_{T,jj}, ..., e_{T,nj})| \qquad (3.5)$$

where $MDE_{T,i}^{CD}$ is the Maximum Difference of Errors when the resolution of CPs is constant and $MDE_{T,j}^{CP}$ is the Maximum Difference of Errors when the resolution of CDs is constant, for a return period T. Following this approach, we can investigate the dominant source of uncertainty (i.e., CDs or CPs) in the system. Also, this can potentially help verify if using the combination of lower resolution CPs and higher resolution CDs could reduce the computational demand, while maintaining a high level of detail in the simulations.

3.3   Results

This section is structured as follows: in section 3.3.1, mean annual hydrographs of simulations are presented. Section 3.3.2 provides the results regarding spatial distribution of hydrological variables. Section 3.3.3 gives the results related to the model efficiency (KGE of simulation) and corresponding uncertainties. Section 3.3.4 provides analyses regarding the uncertainties of extreme flows.

3.3.1   Annual Hydrographs

Figures 3.3 and 3.4 display the mean annual cycle of simulated and observed streamflows for 3- and 24-hour time steps. As discussed in section 3.2.4, for each catchment and model, 16 simulations are available, which is the combination of 4 sets of CP s and 4 CDs resolutions. The figures show the entire period of calibration and validation. Furthermore, winter streamflow has been removed for the 3-hour time step due to a lack of observation data. The results are presented according to the catchment area: the top row shows larger catchments (> 1000 km) whereas the bottom row shows smaller catchments (< 500 km). In Figure 3.3, WaSiM is used to simulate streamflow. The uncertainty bounds in the figures demonstrate the sensitivity of the model to variations of the spatial resolution. Such uncertainty can be found in most of the cases, regardless of the catchment size and time step (3 hours or 24 hours). The Croche, Aux Brochets,

and Boyer catchments, which show notable uncertainties, belong to the groups of large, medium and small size catchments, respectively. Thus, no clear link between the size of the catchment and the degree of uncertainty can be found in this study (hypothesis i). By contrast, the impact of the time step on the uncertainty can be observed for the catchments mentioned above, as the simulations with a 3-hour time step show wider uncertainty bounds (hypothesis ii).
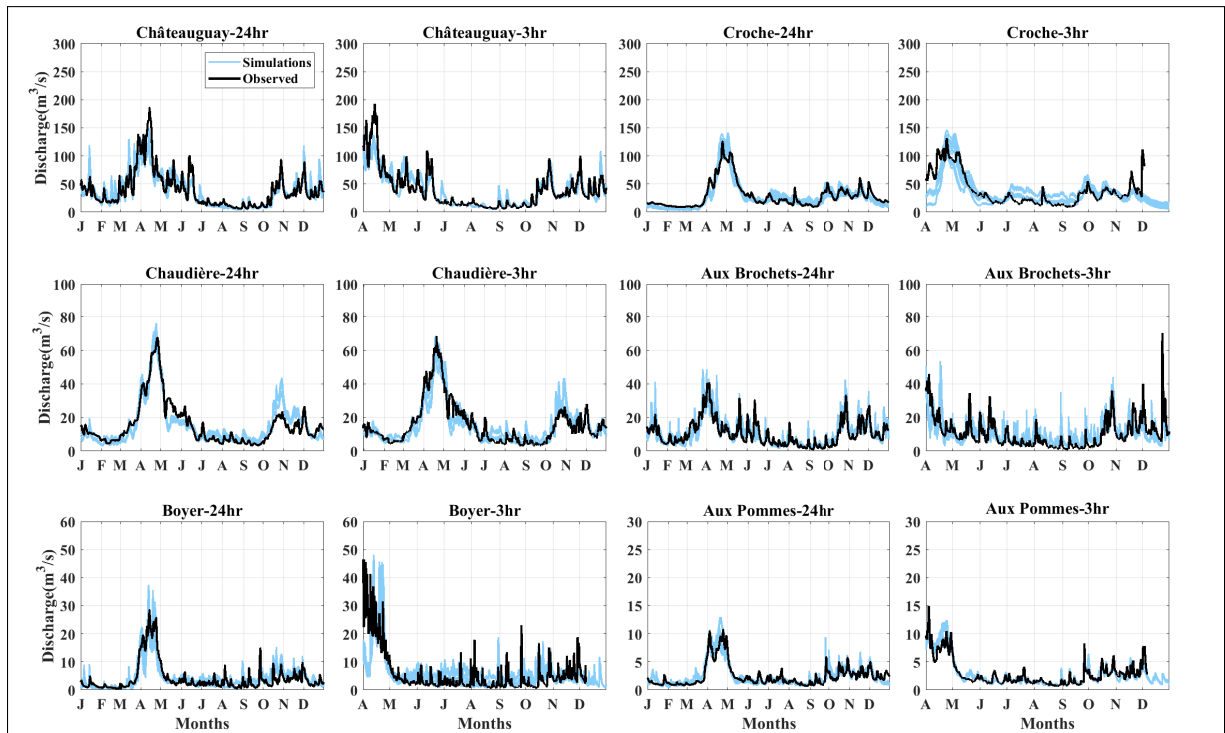
Figure 3.3   Annual hydrographs of the selected catchments simulated by WaSiM and compared to observed data. The modeling time-steps are 24 and 3 hours. The responses are arranged according to the size of the catchments: large catchments (> 1000 km) are on the top row; medium catchments (between 500 and 1000 km) are on the middle row; large catchments (< 500 km) on the bottom row

Figure 3.4 shows the Hydrotel simulations, when the number of HRUs are kept constant (Hydrotel1, see section 3.2.4). Compared to the WaSiM simulations, the model shows less sensitivity to a changing spatial resolution (hypothesis iii). The only exception is the Aux Pommes catchment, in which a large disparity between simulations can be observed. Furthermore, the uncertainty bound is visible for the Croche catchments. Regarding the impact of time steps,

unlike WaSiM, no systematic pattern emerged. For simulations with Hydrotel2, a slight widening of the uncertainty bounds can be observed (Figure I-1 in the supplementary material).
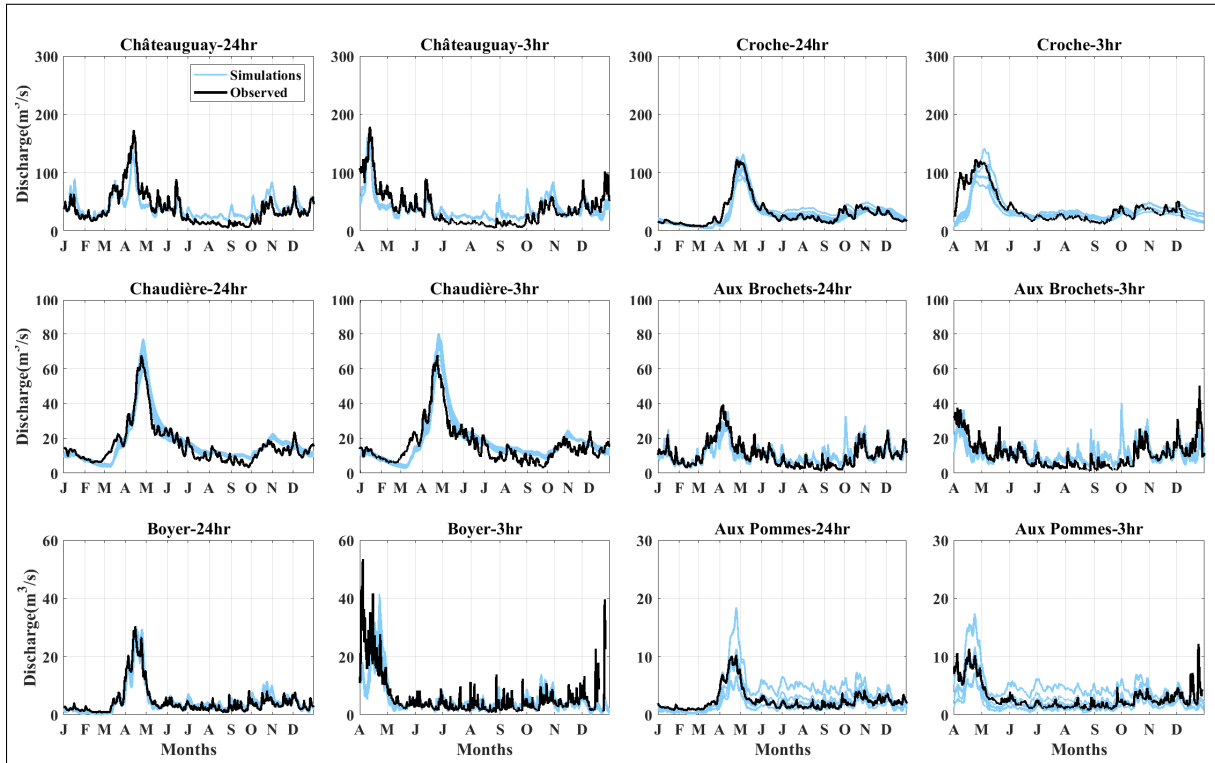
Figure 3.4   Annual hydrographs of the selected catchments simulated by Hydrotel (Hydrotel1) and compared to observed data. The modeling time-steps are 24 and 3 hours. The responses are arranged according to the size of the catchments: large catchments (> 1000 km) are on the top row; medium catchments (between 500 and 1000 km) are on the middle row; large catchments (< 500 km) on the bottom row

Figure 3.5 shows the variance of the ensemble of simulations corresponding with different streamflow quantiles per catchment and time-step. Having 6 variances (one per catchment), we then calculated the mean and standard deviation (STD) of variances to facilitate the comparisons between models. These values are located in the 7th and 8th rows of each subplot. According to Figures 3.3 and 3.4, the results for catchments Aux Pommes simulated by Hydrotel and Boyer simulated by WaSiM suggest that there are inconsistencies between parameters and catchment descriptors for simulations with 100 m resolution, which results in poor efficiencies. Therefore, we excluded the poor simulations for catchment Aux Pommes-Hydrotel (i.e. all combinations of

CP_{100} and CD_{100} except $CP_{100}CD_{100}$). In addition, for Boyer-WaSiM we removed simulations $CP_{100}CD_{250}$, $CP_{100}CD_{500}$, $CP_{1000}CD_{100}$ (the analyses regarding the poor simulations of the Aux Pommes and Boyer catchments can be found in Section 3.4, Figures 3.15, and I-9).
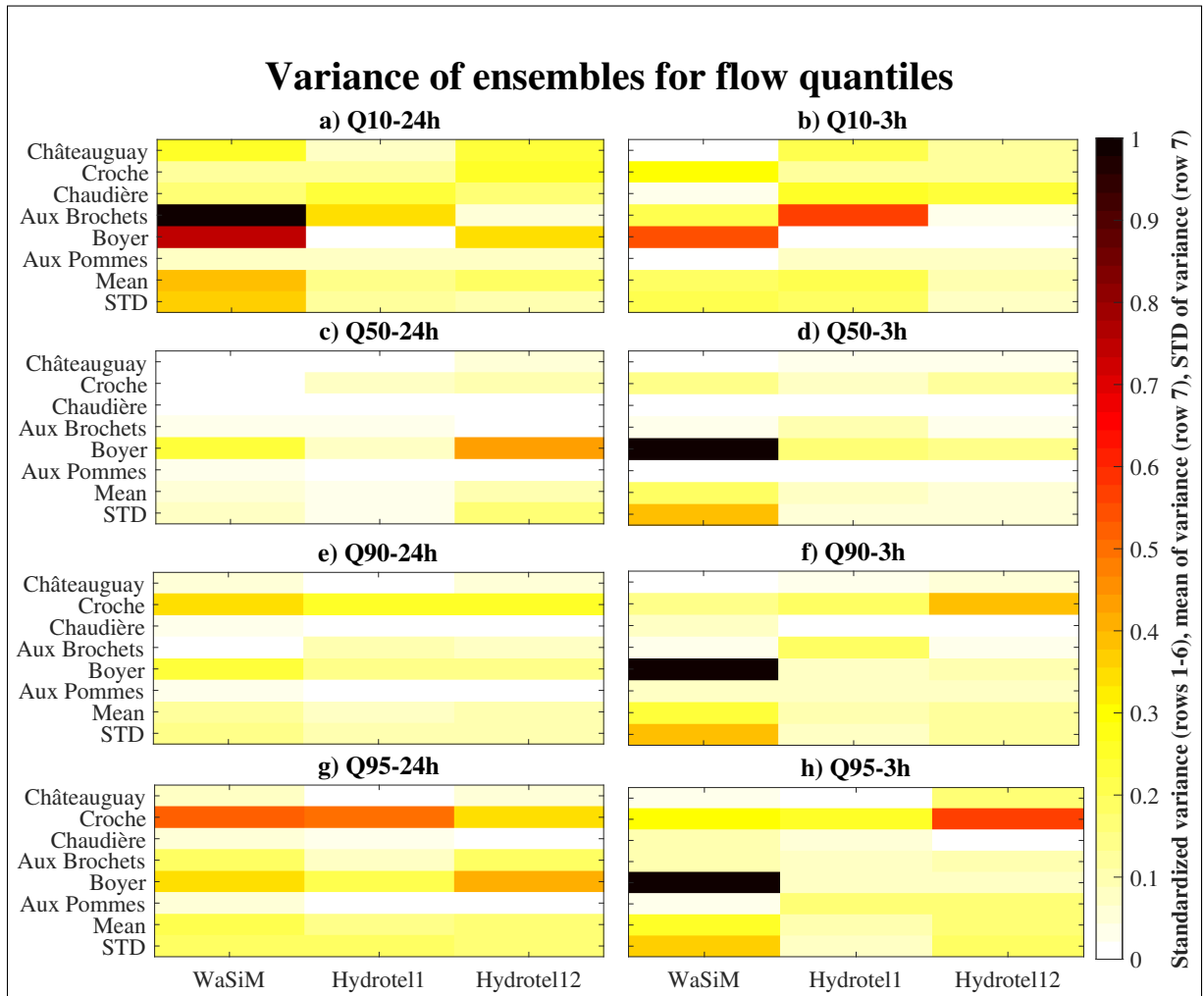
Figure 3.5   Standardized variance of ensemble of flow quantiles (Q10, Q50, Q90, and Q95). For each subplot: Row 1-6 is the standardized variance of quantiles. Row 7 is the mean and row 8 is the STD of the variances over all catchments

Regarding Hypothesis i (i.e. catchments size), these results show that the size of a catchment does not play an important role in determining the uncertainty related to the level of spatial discretization. Looking at *g* for example, such uncertainty can be seen for Croche (a large catchment), Aux Brochets (a medium catchment), and Boyer (a small catchment). Regarding

Hypothesis ii (i.e. time-step of simulations), no pattern can be observed for Hydrotel1 and Hydrotel2 comparing the results for 3- and 24-hour time-step. However, the variance of quantiles for the 3-hour time-step is higher than that of 24-hour for WaSiM. A portion of this difference might be explained by a lack of, or noisy streamflow observations for the 3-hour time-step (e.g. catchment Boyer). We performed a Student t-test and a Wilcoxon rank sum test to find any significant change in the variances when the time-step changes. The results of the tests show no statically significant change. Finally, Figure 3.5 shows that the variance of the simulations is higher for WaSiM than for Hydrotel2 and Hydrotel1 for most cases. In addition, Hydrotel1 shows the lowest variance among the models. These observations support the validity of Hypothesis iii.

3.3.2 Spatial distribution of the hydrological variables

Figures 3.6 and 3.7 demonstrate the effects of different levels of spatial discretization on the spatial distribution of actual evapotranspiration (AET), snow depth, and baseflow across catchment Châteauguay simulated by WaSiM and Hydrotel1 respectively. Here we selected four simulations for which catchment descriptors (CDs) and calibration parameters (CPs) resolutions are identical (i.e. $CP_{100}CD_{100}$, $CP_{250}CD_{250}$, $CP_{500}CD_{500}$, $CP_{1000}CD_{1000}$- the numbers represent resolution in m) to highlight the impact of spatial discretization. In Figure 3.6, variability between the simulations with different spatial resolutions is visible for all variables. For example, the distribution of $CP_{100}CD_{100}$ for baseflow (3rd row) shows a variability across the catchment, which is reduced by increasing the scale. This is expected, as lowering the spatial resolution smooths the topography of the catchment in the model, resulting in a more uniform hydrologic response.

For Hydrotel1 simulations in Figure 3.7, the spatial variability across scales is less than that of WaSiM. Considering the example of baseflow, a similar pattern of spatial distribution can be seen for all simulations, regardless of the resolution. Figures 3.6 and 3.7 confirm the findings of section 4.3.1, in which WaSiM shows a higher degree of sensitivity to changes of scale. These results are in accordance with the third hypothesis. Further analyses regarding the spatial distribution of AET can be found in supplemantery material (Figures I-2 and I-3).
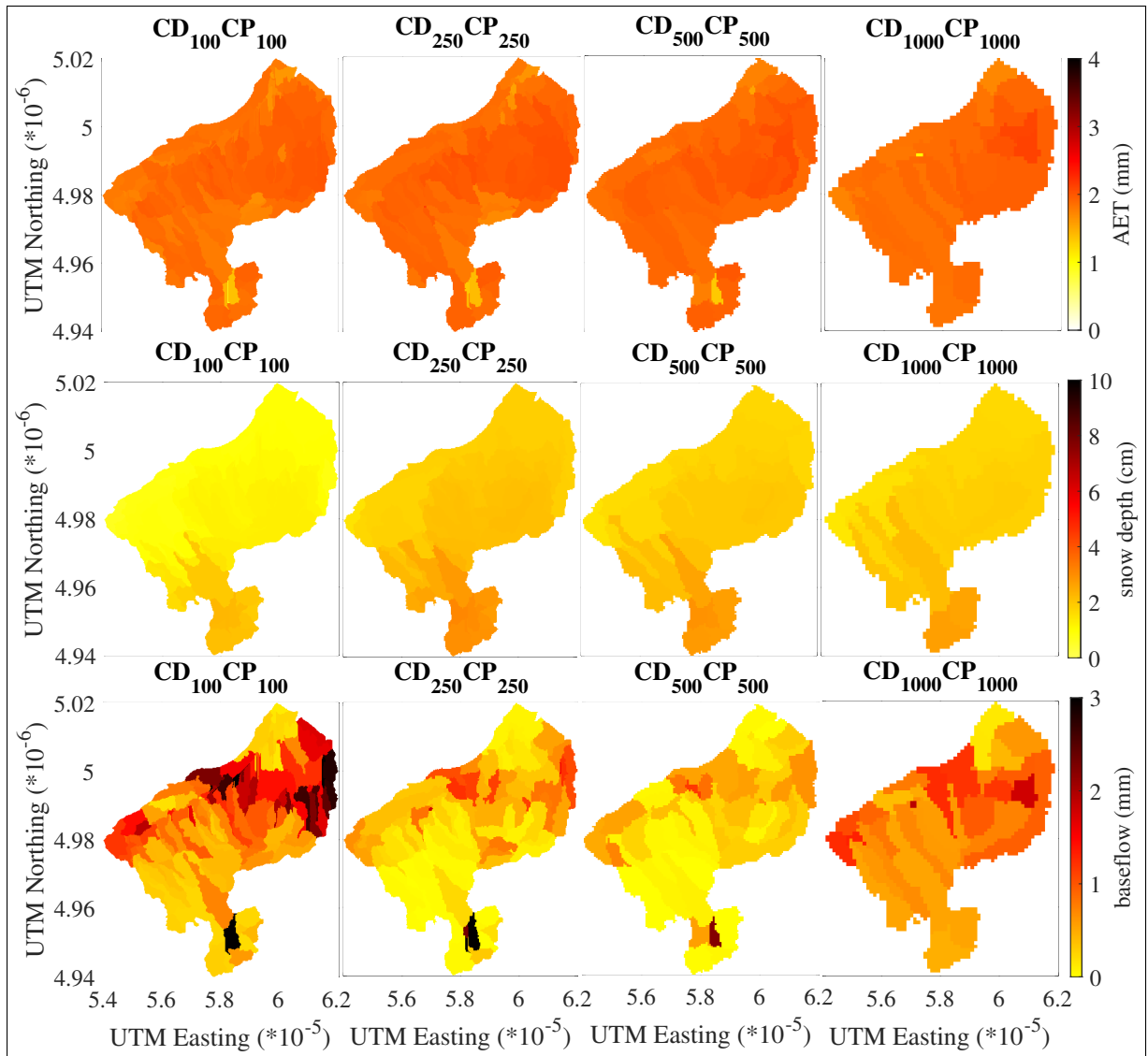
Figure 3.6 Spatial distribution of hydrological variables across different resolutions for catchment Châteauguay simulated by WaSiM. First row: actual evapotranspitation (AET (mm)); second row: snow depth (cm); and third row: baseflow (mm)
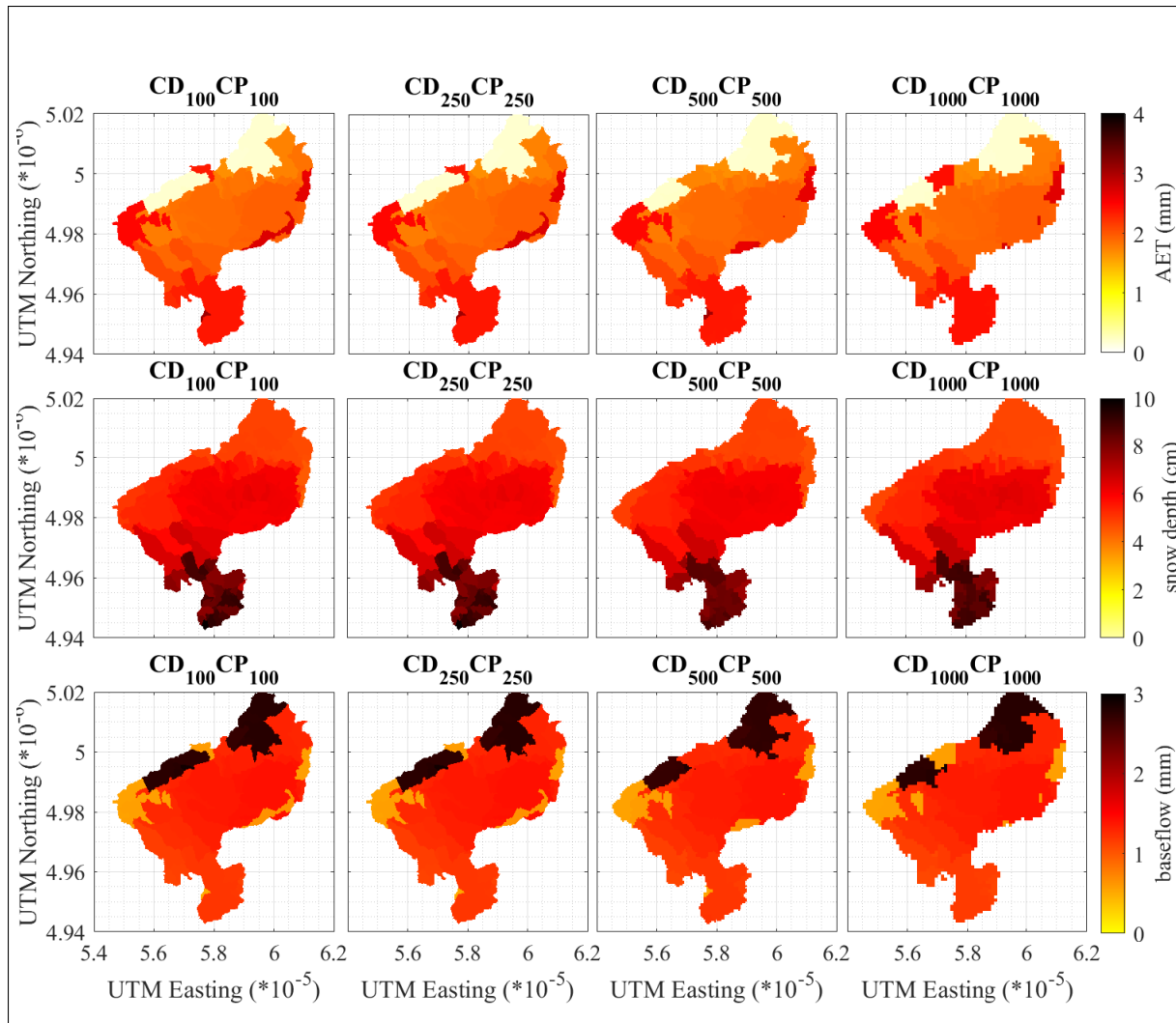
Figure 3.7   Spatial distribution of hydrological variables across different resolutions for catchment Châteauguay simulated by Hydrotel1. First row: actual evapotranspitation (AET (mm)); second row: snow depth (cm); and third row: baseflow (mm)

3.3.3   General performance of the simulations

Figures 3.8 to 3.10 illustrate the performance of the simulations through calibration and validation periods for the catchments, according to the Kling Gupta criterion. Here, we represent how the efficiency of simulations varies following two sources of variability: a primary source, which is caused by direct changes to the catchment descriptors; and a secondary source, which is caused

by any change in the calibration parameters. However, the latter is itself caused by changing the resolution of CDs. We assign a marker and a color to each simulation, which represent the resolution of CDs and CPs respectively.

Figure 3.8 demonstrates the performance of the simulations by WaSiM. Although the number of optimization trials is limited (150) due to the intensive computational demand of the model, the efficiency is high (> 0.8) for most cases. Furthermore, as R^2 values show, the model demonstrate a robust performance for both the validation and calibration periods. It is notable that, except for the Châteauguay and Chaudière catchments, the variability of KGE values is visible, as a result of changes in resolution. In addition, no systematic pattern regarding the relationship between catchment size and uncertainty can be identified. Interestingly, the maximum spread can be seen in Boyer catchment, which is small (191 km^2). In terms of temporal resolution, for most of catchments, the simulations with a 3 hour time step display a slightly higher variability than those with a 24-hour time step (Table I-1 in the supplementary material shows that this change is not significant for most catchments).
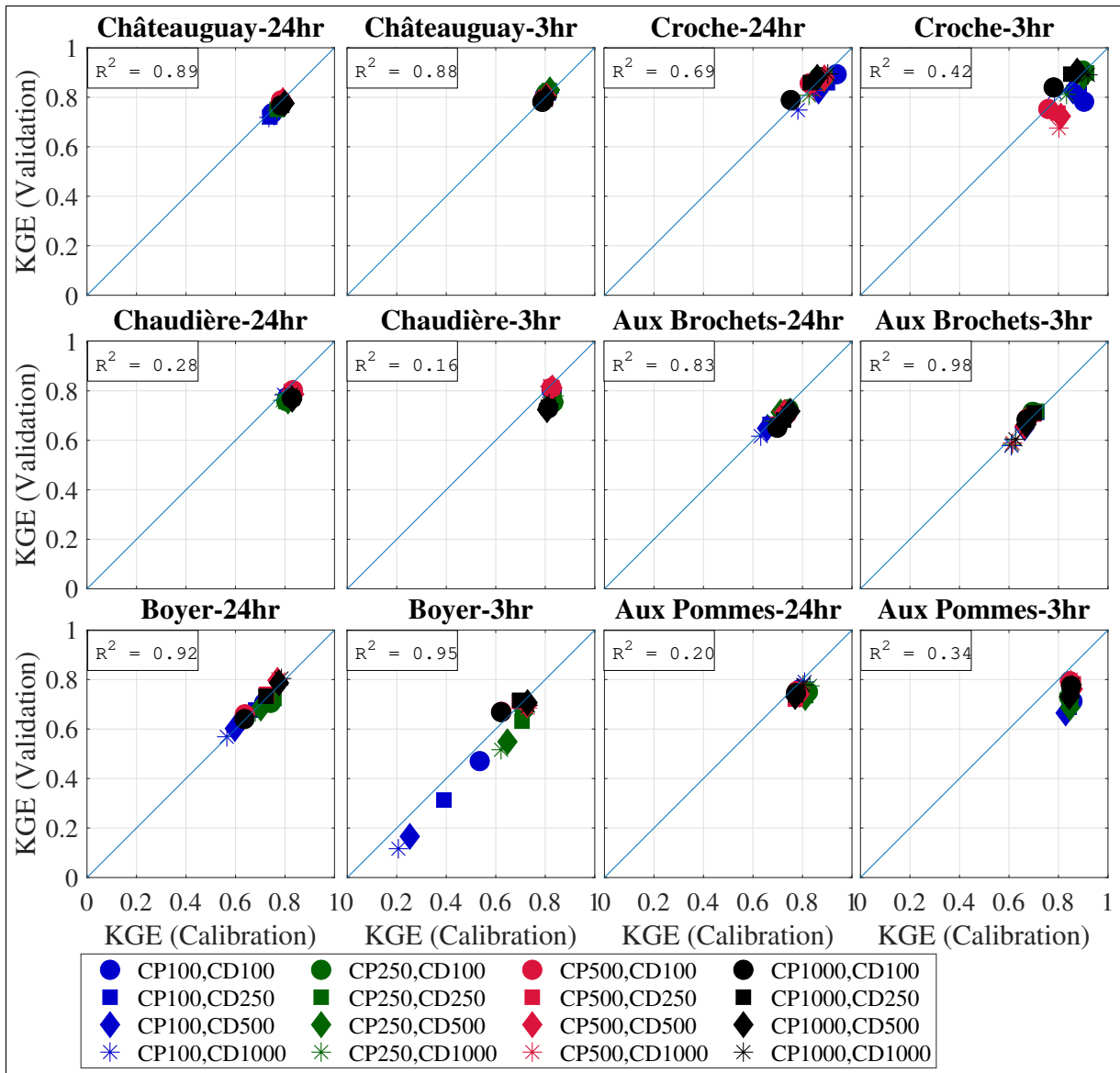
Figure 3.8   Efficiency of WaSiM in reproducing streamflow for the calibration and validation periods. Here, CP and CD represent calibration parameters and catchment descriptors respectively and the numbers assigned show the resolution in *m*

Figures 3.9 and 3.10 show the KGE of simulations by the Hydrotel1 and Hydrotel2 configurations, where 500 optimization trials have been used for each case. In general, the efficiency of simulations with Hydrotel is lower than with WaSiM (> 0.7), even though the number of optimization trials for Hydrotel exceeds those of WaSiM. From another standpoint, in practice,

if the computational power and time are limited, then a semi distributed model such as Hydrotel which has a lower sensitivity to variations in spatial resolutions can be a good choice for simulations (hypothesis iii). Nonetheless, according to R^2 values, the models demonstrate a robust performance for the calibration and validation periods. Furthermore, the Aux Pommes catchment depicts a large variability in the spread of the simulations. Figures 3.9 and 3.10 reveal that a major drop in the performance often occurs when the highest resolution (100 m) of CP (or CD) is combined with the lower resolution of CD (or CP, i.e. 100, 250, 500, 1000 m). Remarkably, such a pattern holds for the WaSiM simulation of the Boyer catchment with a 3-hour time step in Figure 3.8, where a major decline in KGE is seen in simulations (blue). This highlights the issue of compatibility between the resolution at which parameters are calibrated and the resolution at which the model is simulated (This will be further explained in Section 3.4 and in the supplementary material). Comparing Figures 3.10 and 3.9, it can be seen that the spread of the simulations is higher for Hydrotel2 than for Hydrotel1. This is an expected outcome given the scheme used for Hydrotel2, in which the numbers of HRUs are changed in accordance with the resolution of CDs .
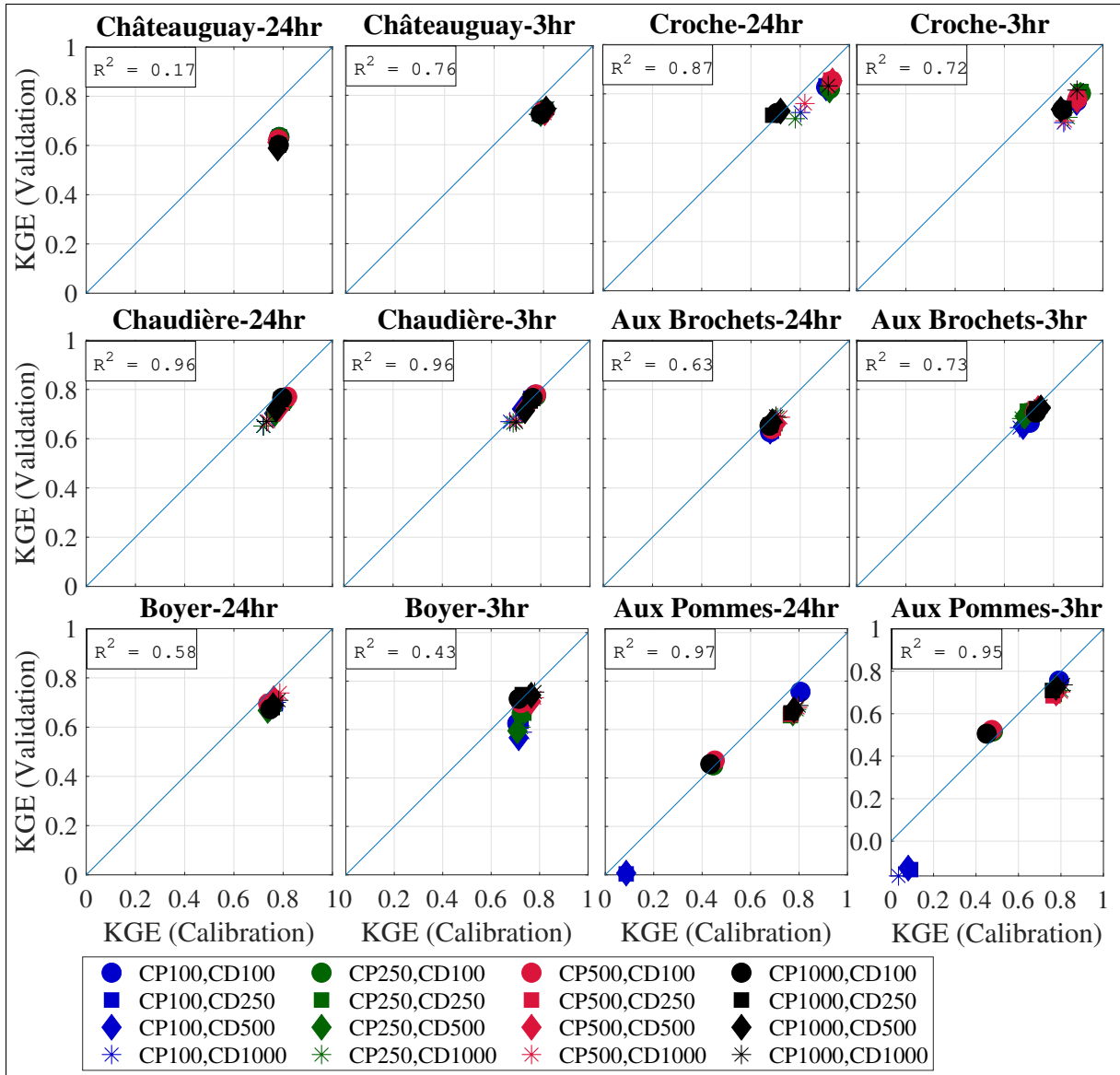
Figure 3.9 Efficiency of Hydrotel (Hydrotel1) in reproducing streamflow for the calibration and validation periods. Here, CP and CD represent calibration parameters and catchment descriptors respectively and the numbers assigned show the resolution in *m*
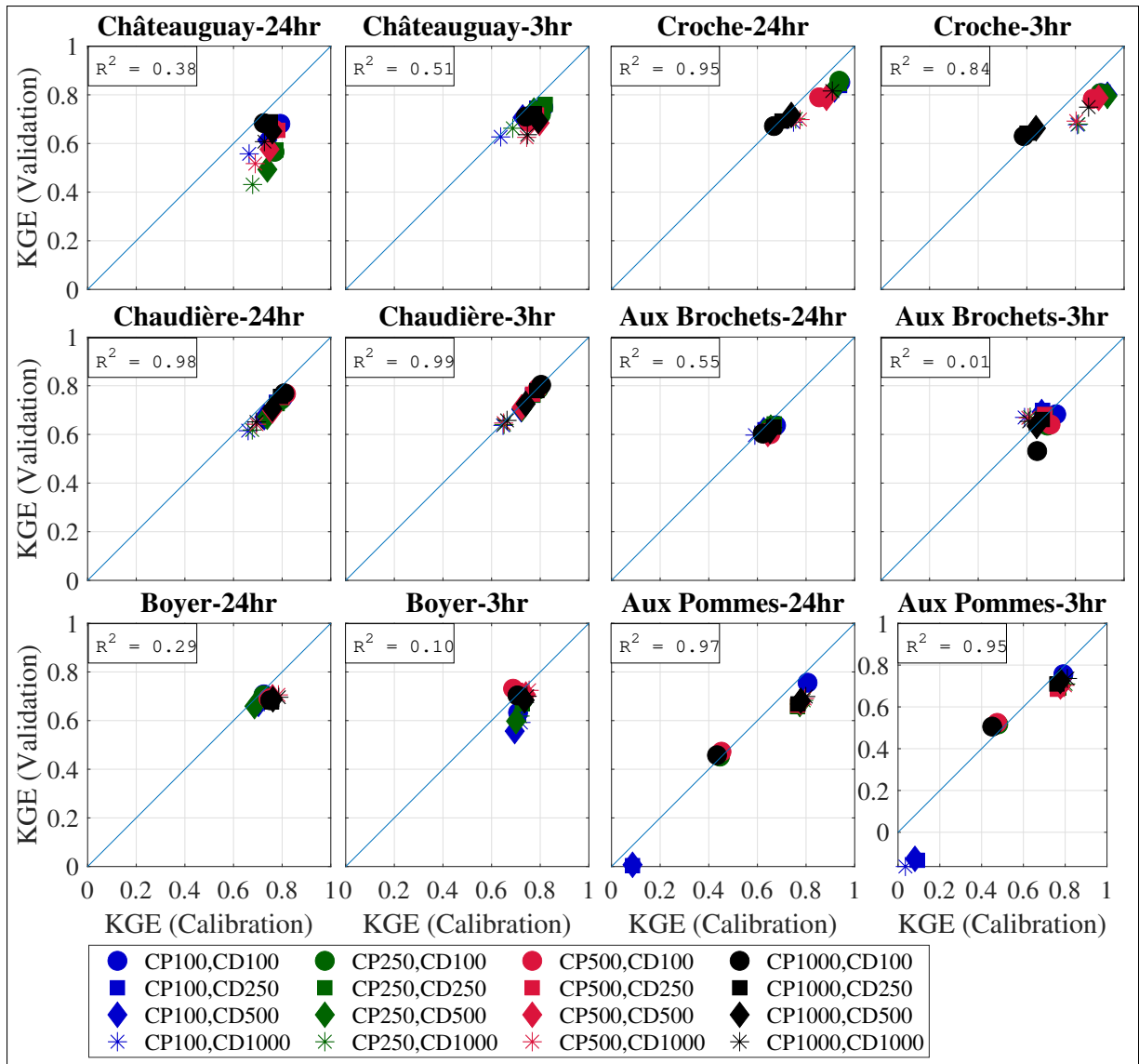
Figure 3.10 The efficiency of Hydrotel (Hydrotel2) in reproducing streamflow for the calibration and validation periods. Here, CP and CD represent calibration parameters and catchment descriptors respectively and the numbers assigned show the resolution in *m*

Looking at Figures 3.8 to 3.10, no systematic pattern can be detected in terms of the impact of variabilities corresponding to CDs or CPs. In some cases, the CDs are dominant (the markers grouped together), while in others, CPs are dominant (colors grouped together) and for the rest of the cases there is no clear pattern (hypothesis iv). The figures, however, reveal that the best performance is not necessarily correlated with the highest possible resolutions of CDs

and CPs. Indeed, the combinations of the lowest resolutions (CP_{1000} and CD_{1000}), which are shown by black colors and asterisk shape markers, are among the top performing simulations. This is important for practical applications, as using a combination of lower resolution CDs for calibration and high resolution CDs for simulation could substantially reduce the computational costs while maintaining the detail of simulations.

3.3.4   Uncertainty of extreme streamflows

Figures 3.11 to 3.13 show the relative error when the models are used to simulate floods with 5-, 10-, 20-, and 50-year return periods. We fitted the Log-Pearson distribution to the annual maxima of the simulated and observed streamflows for the 2000-2017 period and extracted the flood events corresponding to the return periods mentioned above. The spread of the boxplots show the difference in relative error (Equations 3.3) of all simulations (i.e., for the ensemble of 16, which is combination of CDs and CPs in each case) generated by changes in spatial resolution. Given the nature of extreme events, which comprise streamflows with large magnitudes, the noticeable spread of simulations shown in these figures highlights the importance of spatial discretization for flood modeling. Figure 3.11 demonstrates the relative error of extreme events simulated by WaSiM. In agreement with the previous observations, a spread can be detected across different catchment sizes, (i.e. Croche, Aux Brochets, Aux Pommes) and a systematic relationship between extreme flow and catchment size cannot be identified (hypothesis i). Moreover, there is no significant relationship between the spread and the time step of the simulations.
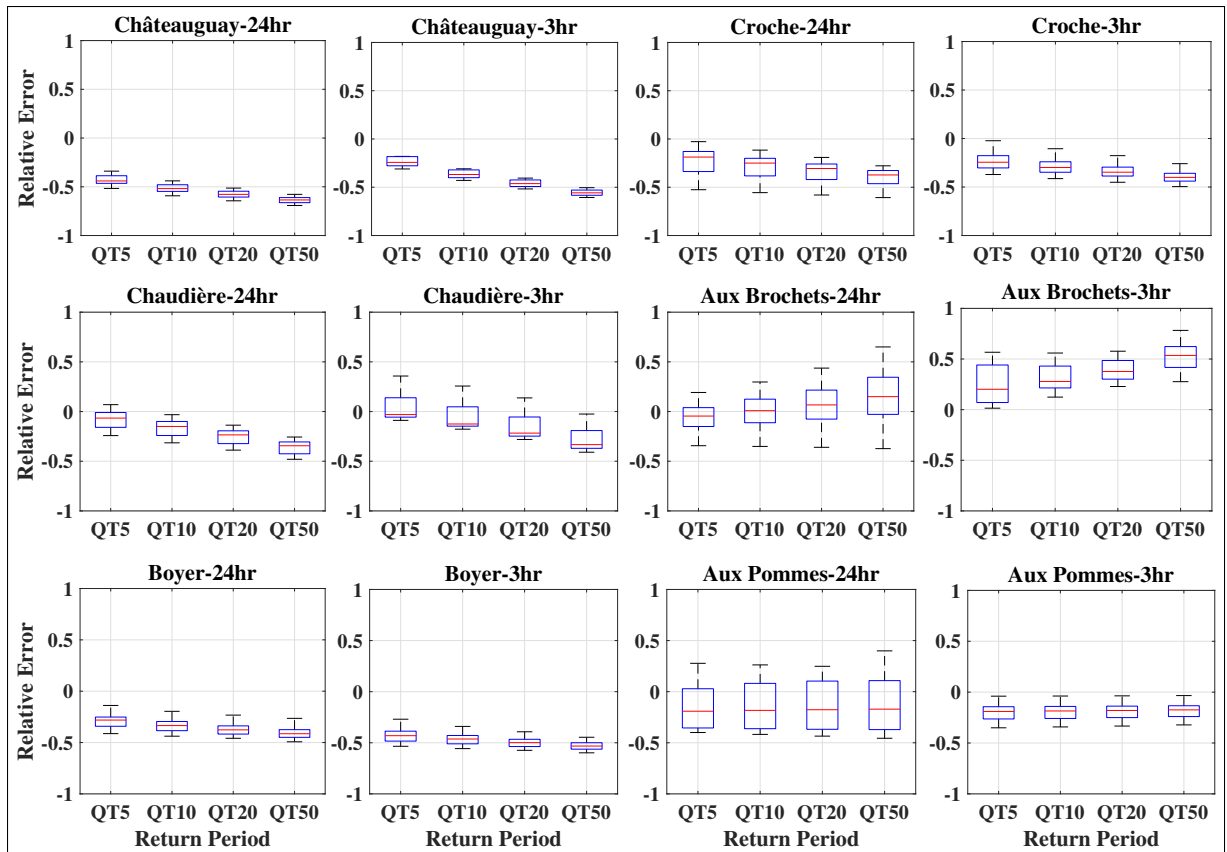
Figure 3.11   Relative error for the simulation of summer-fall floods with 5-, 10-, 20-, and 50-year return periods using WaSiM. QT represents a flood event with the specific return periods

Figures 3.12 and 3.13 show the relative error of flood simulations produced by the Hydrotel1 and Hydrotel2 configurations. The response of Hydrotel1 to extreme flow is similar to other figures (i.e. annual hydrographs and KGE) discussed earlier. While the magnitude of error is higher as compared to WaSiM, the model shows a smaller spread of relative errors (hypothesis iii). Nonetheless, the spread of relative error is visible across different catchment sizes (Châteauguay, Aux Brochets, and Aux Pommes), which refutes the possibility of a correlation between the catchment size and the uncertainty of extreme flow (hypothesis i). However, the time step chosen for the simulation is important, as the width of the boxplots corresponding to the 3-hour time step is larger than for the 24-hour time step (hypothesis ii). Simulations with Hydrotel2 exhibit a noticeably larger uncertainty for extreme streamflows as compared to Hydrotel1, particularly for

the Châteauguay and Aux Brochets catchments. This is in line with the earlier observations discussed in Figures 3.9 and 3.10, where the uncertainties corresponding to Hydrotel2 are higher than for Hydrotel1 due to the change in the numbers of HRUs for Hydrotel2. Finally, considering Figures 3.11 to 3.13, the return period does not appear to influence the uncertainty of the simulations. Indeed, the spread of the simulations for different return periods is similar, per catchment.
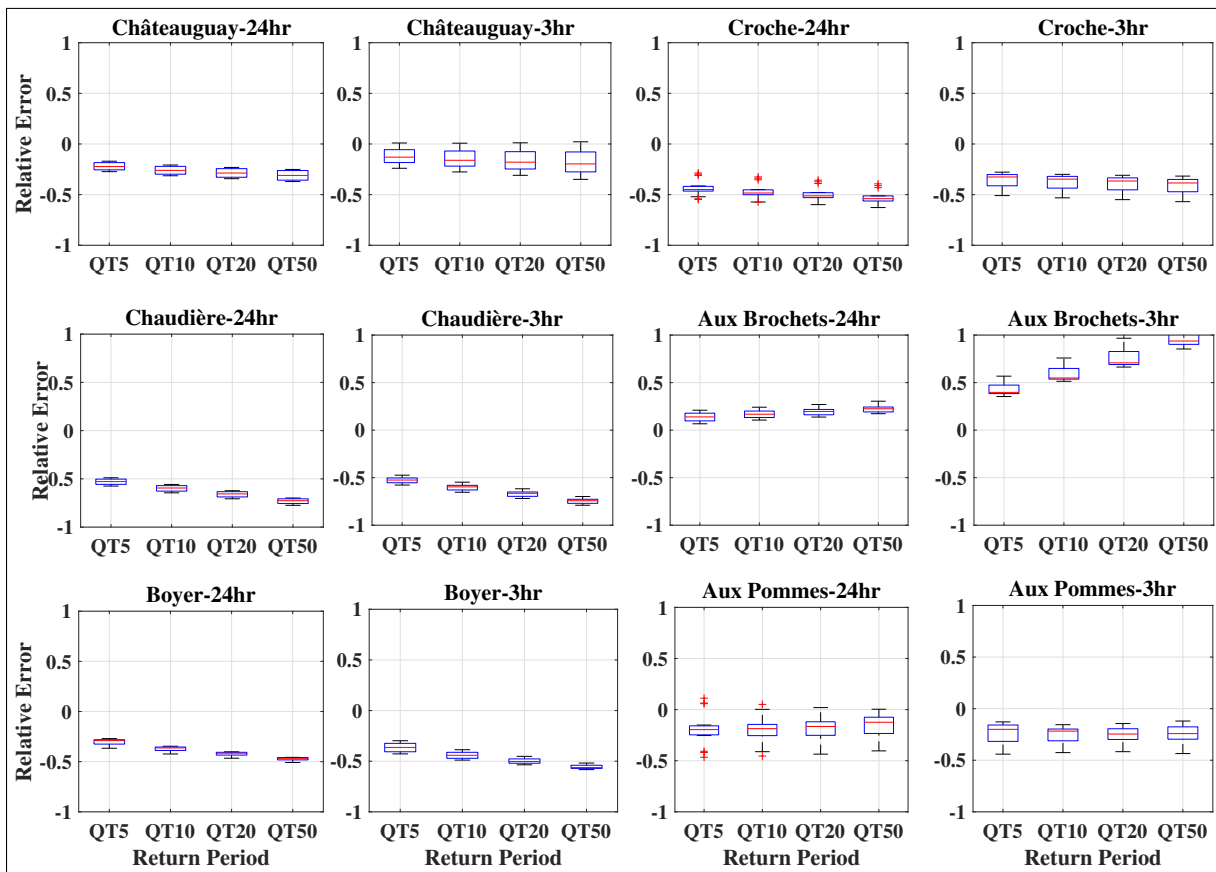
Figure 3.12   Relative error for the simulation of summer-fall floods with 5-, 10-, 20-, and 50-year return periods using the Hydrotel1 configuration. QT represents a flood event with a specific return period
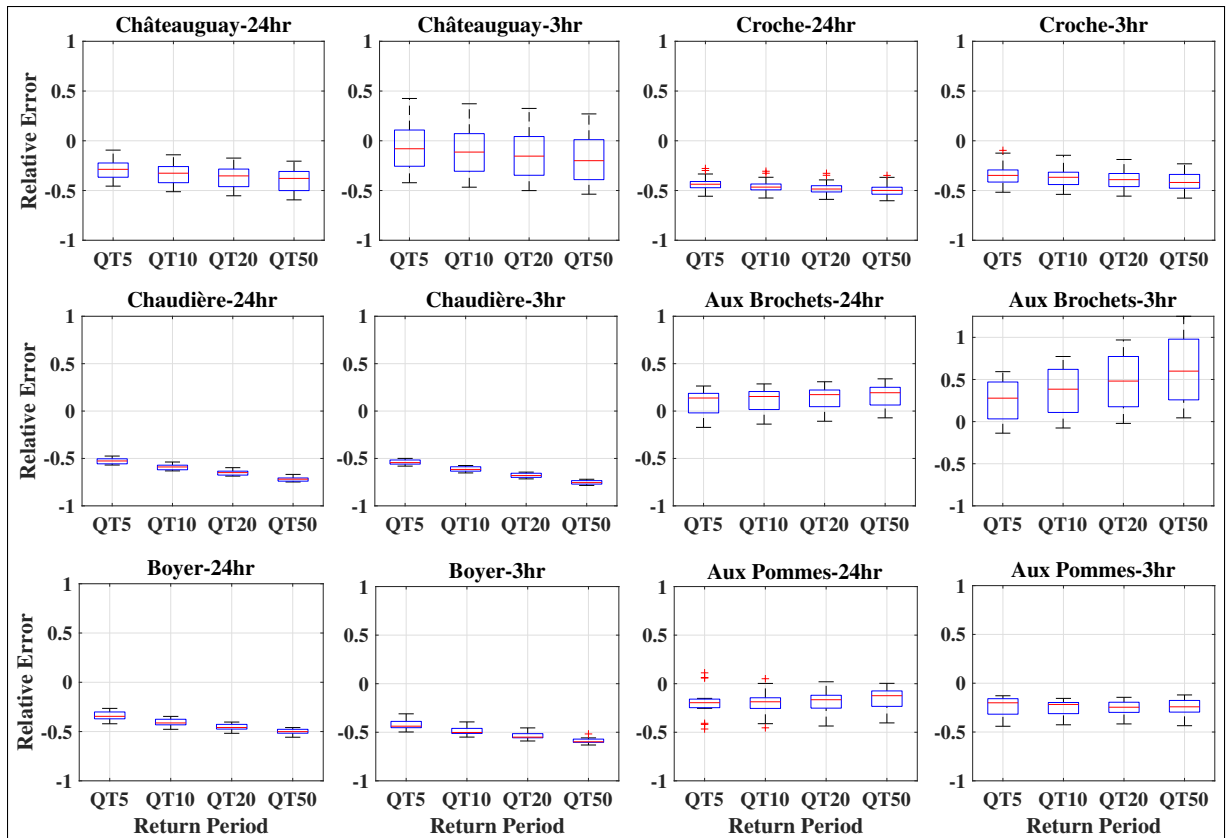
Figure 3.13   Relative error for the simulation of summer-fall floods with 5-, 10-, 20-, and 50-year return periods using the Hydrotel2 configuration. QT represents a flood event with a specific return period

Figures 3.14, I-4, and I-4 illustrate a separation of the total uncertainty for extreme streamflows into contributions from CDs and CPs. The separation procedure is carried out following section 3.2.4. In these figures, RN represents the resolution of simulations and QTN represents the return period. The vertical and horizontal axes are the Maximum Difference of relative Errors (MDE) of CDs and CPs respectively, as defined in Equations 3.4 and 3.5.

Figure 3.14 depicts the results of simulations with WaSiM. For most catchments, the contribution of CPs to the total uncertainty is larger than that of CDs. For instance, the MDE of CPs in Châteauguay is between 0.1 to 0.2, while the MDE of CDs is around zero. The same pattern can be seen for Croche, Chaudière, Aux Pommes (3 hour), and Boyer (3 hour). This, however, is not the case for all catchments. For the Aux Brochets (3 and 24 hour) and Aux Pommes (24 hour)

catchments the MDE corresponding to CDs is equal to or larger than that of CPs. The dominance of MDE of CDs is evident, particularly for Aux Brochets (3 hour). The reasons for this behaviour are explained in details in the supplementary material (Figures I-6-I-8). Interestingly, the Aux Brochets (24 hours and 3 hours) and Aux Pommes (24 hour) catchments demonstrate the highest range of uncertainty among all catchments. This highlights the importance of accounting for the contribution of CDs to the total uncertainty of extreme streamflow simulations when dealing with catchments that are sensitive to changes in resolution.
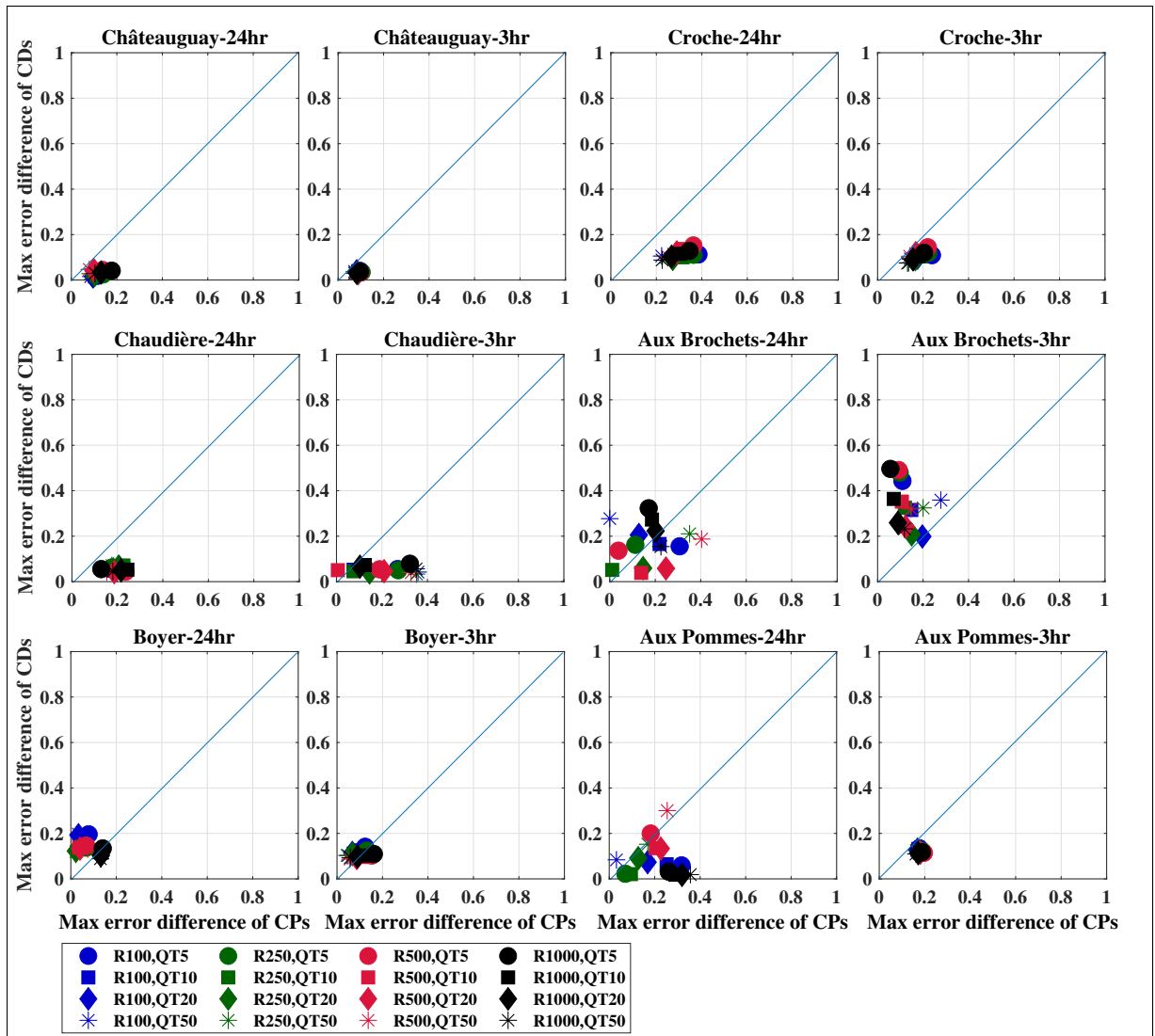
Figure 3.14   Relative error for the simulation of summer-fall floods with 5-, 10-, 20-, and 50-year return periods using WaSiM. QT represents a flood with a specific return period. For instance, QT5 is the flood magnitude corresponding to a 5-year return period. R represents the resolution of CDs or CPs, in which the Maximum Error Difference (MDE) is calculated

3.4   Discussion

In this paper we proposed a novel framework to determine the uncertainty corresponding with catchment heterogeneity. The main objective of this paper is to determine the extent to which the

representation of spatial scale in the modeling can affect the simulations of hydrological variables when the catchment size, time-step of simulation, and the hydrological model vary. Lack of an applicable framework to measure uncertainty of heterogeneity for computationally intensive physically-based hydrological models was a major challenge. A normal procedure to account for parametric uncertainty for example, is to sample from the acceptable range of model parameters, perform multiple simulations, and accept the parameter sets that satisfy a certain criterion (e.g. $KGE > 0.7$, Beven & Freer, 2001). Evidently, applying such procedure is infeasible for the case of catchment heterogeneity. Moreover, the intensive computation of hydrological models limits the implementation of multiple simulations with different resolutions. This issue was dealt with by combining the parameters obtained per resolution with all available CD resolutions and implementing the models per combination to create an ensemble of simulations as discussed in Section 3.2.

To reduce the degree of incompatibility between scale of calibration parameters and catchment descriptors, we maintained the maximum possible similarity between the drainage system structure (stream network, routing channels, sub-catchments etc.) through modifying DEMs and sensitivity analyses of the parameters of the topographic analysis softwares that create the setup required for implementing WaSiM and Hydrotel respectively. We found that for the majority of catchments, such a combination of CDs and CPs results in an acceptable efficiency. In other words, the parameter sets solved for one scale could also be a potential candidate for simulations with another scale. However, this should be said with caution, as mixing the scales might result in poor efficiencies like those observed for Aux Pommes-Hydrotel and Boyer-WaSiM in Figures 3.8 to 3.10.

We discuss more in depth the behaviour of the Boyer catchment simulated by WaSiM at the 3-hour time-step, for which the ensemble of simulations shows a large uncertainty (Figure 3.3) and the KGE displays variability (Figure 3.8). Note that a similar analysis regarding the response of the Aux Pommes catchment simulated by Hydrotel can be found in the supplementary material (Figure I-9). Figure 3.15 demonstrates the distribution of interflow (subplots a and b) and slope (subplots d and e) for 100 m and 500 m resolutions simulated by WaSiM (3-hour

time-step). Subplot c shows the values of interflow storage and potential evapotranspiration (PET) parameters for different spatio-temporal resolutions. In addition, subplot f shows the annual hydrograph of interflow for the outlet zone focusing on peak flow (mid-March to mid-April).

Those figures indicate that two interconnected factors play a role in determining the model behaviour for catchment Boyer. Comparing the spatial distribution of interflow in subplots *a* and *b* of Figure 5, it appears that the interflow decreases when the spatial resolution is decreased in the model. In addition, subplot *c* (Figure 3.15) shows a near 2-fold increase for the coefficient of interflow storage (3-hour) when the spatial resolution is reduced from 100, to 250, 500, 1000 m. A similar behaviour can be seen for the PET coefficient (3-hour) in subplot c. Although it is difficult to fully explain the interactions between those parameters during the calibration period, it seems that the effect of increasing interflow coefficient is balanced by increasing the PET coefficient. For example, assigning a small value to the PET coefficient increases the soil moisture, resulting in less interflow storage and vice versa. However, for the 24-hour time-step, both parameters remain approximately constant (Note that even for the 24-hour time-step, a correlation between PET and interflow storage coefficients across the scales is visible). One reason for observing such marked variations in parameters for the 3-hour time-step roots in the choice of the Hamon equation for the PET submodel, which is empirical and compatible with daily time-step. A more physics-based option compatible with subdaily resolution is the Penman-Monteith equation, but this requires a wide range of data that is not readily available for this study and for many other sites in the world.

Figure 3.15 *f* demonstrates the mean annual hydrograph of interflow at the outlet of the Boyer catchment, focusing on the peak flow from mid- March to mid- April at the 3-hour time-step. Here, all combinations of simulations with CP_{100} (i.e. $CP_{100}CD_{100}$, $CP_{100}CD_{250}$, $CP_{100}CD_{500}$, $CP_{100}CD_{1000}$) are shown with larger line thickness. These simulations show higher peaks in the ensemble compared to the other simulations. Since all simulations are driven from the same parameter set (CP_{100}), with low interflow storage (i.e. more saturated soil) a higher interflow for those simulations is expected. Interestingly, by increasing the spatial scale of CDs, from

100, to 250, 500, 1000 m, while keeping CP constant (i.e. CP=100), an increase of the peak value of interflow is observable (see the thick lines in subplot f). This behaviour could be attributed to the spatial correlation between the slope of the catchment and the spatial distribution of interflow. Comparing subplots a and b with d and e respectively shows that more marked slopes induce lower interflow values and vice versa. Altering the slope by increasing the spatial scale affects the drainage density of the catchment in the model. In fact, smoothing the slopes by increasing the spatial scale of CDs causes a reduction in the drainage density, resulting in more infiltration and the observed increase in interflow. This leads to an overestimation of streamflow at the outlet of the catchment, as shown in Figure 3. Thus, a lower resolution of CDs combined with CP_{100} (i.e. $CP_{100}CD_{250}$, $CP_{100}CD_{500}$, $CP_{100}CD_{1000}$) overall results in higher interflow values for catchment Boyer, adding to the uncertainty attributable to the choice of resolution for the spatial discretization (Figure 3.3) and the associated variability in KGE values (Figure 3.8).
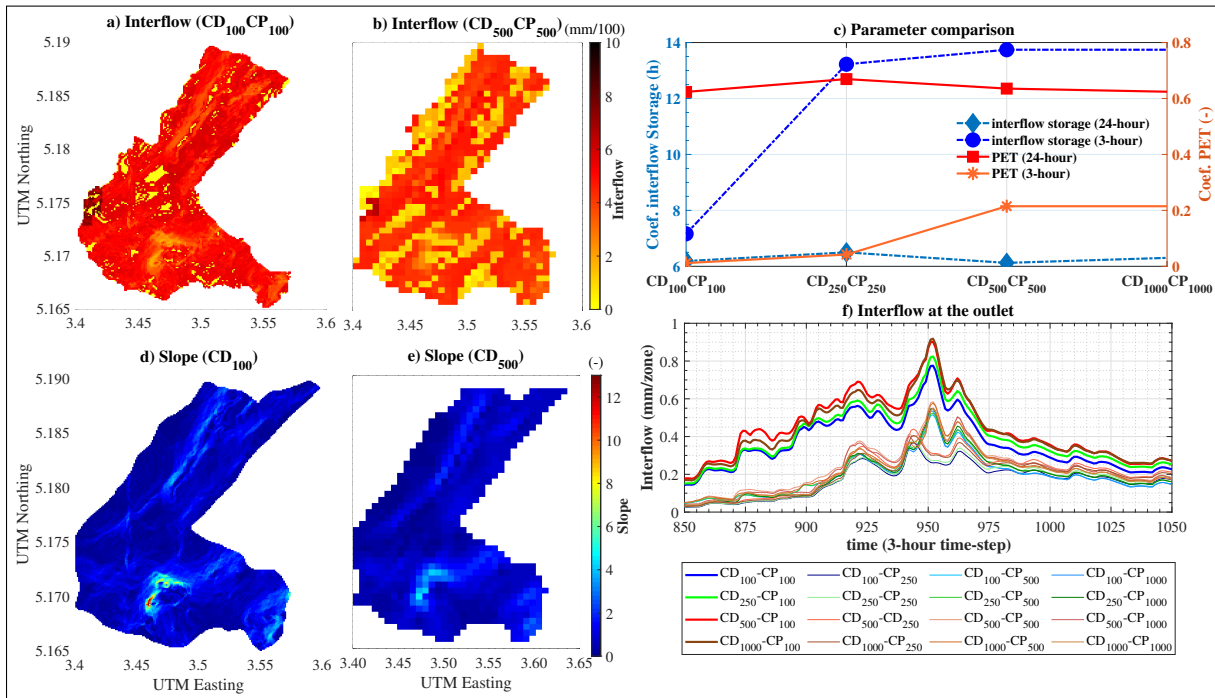
Figure 3.15   Analyses of catchment Boyer simulated by WaSiM at a 3-hour time-step: a and b) distribution of interflow for 100 and 500 m^2 simulations; c) Interflow storage and PET parameters for 3- and 24-hour time-steps; d and e) distribution of slope for 100 and 500 m simulations; f) Annual cycle hydrograph at the outlet of the catchment

In terms of land use and topographic characteristics of the catchments (Table S2 in the supplementary material), we cannot single out one major characteristic that is significantly different from other catchments. A combination of multiple factors might affect the simulations. For instance, agricultural land occupies a major portion of the surface area (66%). In addition, the share of deciduous forest is 14%, which is the least among other catchments. On the one hand, these numbers reveal major human intervention (agriculture and deforestation) throughout the catchment, potentially affecting different hydrological processes that include PET, AET, direct runoff, etc. On the other hand, the catchment generally has a low elevation and slope accompanied with a high drainage density. Such topographic features are sensitive to changes in spatial resolution, as shown for catchment Aux Brochets (Figures I-6-I-8). Catchment Boyer also includes 5% bogland, which is the highest of all catchments, another factor that further complicates its modeling.

As discussed in section 3.1, previous research addressing the uncertainty of heterogeneity by focusing on variation of spatial resolution so far have mostly focused on urban catchments (e.g., Cao *et al.*, 2020a; Krebs *et al.*, 2014; Zhou *et al.*, 2017; Ichiba *et al.*, 2018; Warsta *et al.*, 2017). These catchments are typically very small with a high degree of imperviousness and therefore have different behaviours than natural catchments. However, it is possible to compare results of the present study with those of urban catchments. According to Ichiba *et al.* (2018) and Warsta *et al.* (2017), changing spatio-temporal resolution of the urban distributed models results in over- or under-estimation of streamflow, but a relationship between the direction of such a change and spatial resolution cannot be determined. Similarly, our study showed that no pattern can be detected regarding a general relationship between spatio-temporal representation of the modeling and increase or decrease of streamflow. This can be explained by the fact that, each catchment has a unique distribution of slope and land-use, to which the hydrological processes such as interflow and AET have non-linear and complicated relationships such as what we showed in Figures 3.15, I-6 to I-6. Therefore, it is difficult to estimate the direction of streamflow variation based solely on the spatio-temporal resolution of the modeling.

Out of the limited studies on natural catchments, the findings of Tegegne *et al.* (2019) are partly comparable to our results regarding the Hydrotel configurations with varying numbers of subbasins (i.e. Hydrotel1 and Hydrotel2). Based on their experiments, changing the number of subbasins results in only minor effects on the simulation hydrographs, but it results in significant uncertainty for different flow phases (i.e. different flow quantiles). Another related study in this context by Chen *et al.* (2019b) has reached a similar conclusion that the impact of land-use variation has no major effects on the simulated hydrographs but can change different flow indices (i.e. minimum seven day or maximum daily flow). Similarly, our results show that the spread of streamflow simulations is rather narrow for semi-distributed model in both Hydrotel configurations (Figures 3.4 and I-2). Minor differences in the distribution of other hydrological variables can be observed (Figure 3.7). Moreover, in terms of simulating extreme streamflow, there is a significant amount of uncertainty for WaSiM, Hydrotel1 and Hydrotel2 simulations across all catchments (Figures 3.11, 3.12, and 3.13).

The results from the simulations with the fully-distributed model (WaSiM) are not quite in accordance with the above mentioned conclusions. We showed that using a fully-distributed model with finer time-step leads to a larger spread of simulated hydrographs and significant variation in the spatial distribution of AET, snow depth, and baseflow (Figures 3.3 and 3.6). This demonstrates that the choice of spatio-temporal resolution for fully-distributed models is a key factor in simulating streamflow and other state and flux variables. Therefore, a multi spatio-temporal simulation method is required to account for the corresponding uncertainties. It is worth noting that the aforementioned studies used semi-distributed models and did not resort to multi spatio-temporal simulations, which is the approach used in this paper. Given that, similar researches applying fully-distributed models in the context of assessing the uncertainty linked to catchment heterogeneity are still very scarce.

3.5   Conclusion

We have explored the impact of spatio-temporal discretization to reproduce streamflow and simulate flood events across six catchments located in Quebec (Canada) using two distributed

hydrological models (i.e. WaSiM and Hydrotel including its two different configurations: Hydrotel1, and Hydrotel2). We framed a hypothesis regarding the uncertainty of heterogeneity and broke it down into four main aspects reiterated as follows: Changing the spatial resolution of catchment descriptors generates uncertainty that can potentially impact flood simulations. The catchment area, the modeling time step, and the model structure are the major components used to determine the significance of such uncertainty. Based on the above results and analysis, the following conclusions can be drawn:

1. There is no systematic link between the catchment size and the uncertainty corresponding to the simulation of streamflow, so hypothesis *i* is not verified for our experiment. Regardless of the model used to reproduce streamflow, the uncertainty of heterogeneity has been observed across different catchment sizes (see Figures 3.3, 3.4, I-1 and 3.5 and Table I-1). Interestingly, smaller size catchments (Boyer and Aux Pommes) generate larger uncertainties and larger variabilities in the modeling efficiency (see Figures 3.8 and 3.10), which refutes the assumption that changing the spatial resolution mainly affects larger catchments.

2. The temporal resolution plays only a minor role in the determination of the uncertainty related to spatial resolution, so hypothesis *ii* is also not clearly verified for our experiment. WaSiM and Hydrotel2 showed that a 3-hour time step could moderately increase the uncertainty bounds of simulations for most catchments (see Figures 3.3, 3.5, I-1 , and Table I-1).

3. The model structure is an important driver of the uncertainty related to the spatial resolution of simulations (hypothesis *iii* is verified for our experiment). WaSiM demonstrated a sensitivity to changes in the spatio-temporal resolution of the simulations (See Figures 3.3, 3.4, 3.6, and 3.7). This was expected, given that the model solves Richards Equations for each grid cell, associated with specific catchment descriptors. Hydrotel's conceptualization of infiltration, percolation and groundwater is less physically-based. In its default setting, it cannot adequately capture the uncertainty related to spatial discretization unless change is imposed by altering the number of HRUs (see Figures 3.4 ,3.5 ,I-1, and Table I-1). Moreover, in terms of stability, the results from catchment Aux Pommes simulated by Hydrotel (Figures 3.9, 3.10, and I-9) and Boyer simulated by WaSiM (Figures 3.8 and 3.15)

show that both models have issues for certain combinations of CP and CD, resulting in over or underestimation of streamflow and significant drop in efficiency.

4. Our attempt to separate the total spatio-temporal uncertainty into a portion attributable to CDs and a portion attributable to CPs showed that the latter is the dominant contributor for most of the catchments (hypothesis *iv*-see Figures 3.14, I-4 and I-5). However, there are catchments in which the change of CD resolution is as important or dominant (e.g., the Aux Brochets, Boyer and Aux Pommes catchments in Figures I-6 to I-8, 3.15 and I-9). Based on section 3.4 and the discussion under Figures I-6 to I-9, this might be due to changes in the interactions of hydrological variables once the resolution of simulations is altered (see Figure I-7). Such behavior is expected for relatively flat catchments, but that still includes multiple small hillslopes and valleys (see Table I-2). Indeed, changing the resolution can reduce the impact of an uneven topography, or even eliminate it completely, which can result in an inconsistent hydrologic behaviour and response of the catchment (see Figures 3.15 and I-6 to I-9).

Given the dearth of credible publications addressing the impact of the uncertainty corresponding to the resolution of simulations, many gaps and opportunities remain to be addressed in this line of research. One major area of focus could be the adoption of more advanced physically-based distributed hydrological models to explore the degree of uncertainty, particularly for extreme streamflows. Another focus could be on identifying the key parameters and hydrological processes that are mainly affected by spatio-temporal discretization change. Finally, using a larger set of catchments with different physical characteristics could help provide a better understanding of how they react to variations of the resolution of catchment descriptors. It could also shed light on the importance of accounting for this uncertainty in streamflow simulations and in the assessment of flood events.

CHAPTER 4

MULTI-SCALE FLOOD SIMULATIONS UNDER CLIMATE CHANGE SCENARIOS

Siavash P. Markhali[1] , Annie Poulin[1] , Marie-Amélie Boucher[2]

[1] Department of Construction Engineering, École de technologie supérieure, 1100 Notre-Dame West, Montréal, Québec, Canada H3C 1K3
[2]Civil and Building Engineering Department, Université de Sherbrooke, 2500 Bd de l'Université, Sherbrooke, Québec,Canada J3X 1S1

Section

This study focuses on quantifying the impact of the choice of spatio-temporal resolution and hydrological models on the projection of extreme flow and their link to catchment size. We use two process-based distributed hydrological models forced with a large-ensemble regional climate model (50-member ClimEx dataset) over the 1990-2100 period at different spatio-temporal resolution. The extreme summer-fall flow corresponding with each spatio-temporal resolution was extracted by pooling the members together and computing the empirical cumulative distribution function. The results show that by refining the time-step from daily to sub-daily, the summer-fall extreme flow projected over the future period exceeds that of the reference period for the small but not large catchments. By increasing the catchment size, the hydrological model's contribution to the variability of extreme flow increases. Moreover, the choice of spatial resolution affects the extreme flow trend in terms of magnitude, significance, and direction. But no pattern regarding the catchment size and spatial discretization variations exists.

4.1   Introduction

Flood hazard continues to threaten human life and inflict costs on infrastructures and urban areas, as multiple devastating events have been reported in recent years around the world (Merz *et al.*, 2021). Accurate flood estimation remains a critical issue and the traditional stationary

assumption employed by flood estimation methods, whether empirical or process-based, fails to account for changing climate signal, leading to inaccurate estimations of exceeding probability of peak flow (Blöschl *et al.*, 2013; François, Schlef, Wi & Brown, 2019; Montanari & Koutsoyiannis, 2014). Moreover, a lack of knowledge regarding flood-generating processes at different scales with complex and non-linear catchment responses in space and time complicates the estimation of flood return period using process-based hydrological models (Beven, 2019; Blöschl, 2022b). The present research aims to investigate how the discrete representation of catchments in process-based distributed hydrological models can affect flood projection under climate change scenarios. The study is conducted for snow-dominated Nordic catchments located in Canada.

Global warming is expected to increase the magnitude and frequency of extreme precipitation across different parts of the world (Min, Zhang, Zwiers & Hegerl, 2011; Westra, Alexander & Zwiers, 2013; Alexander *et al.*, 2006; Donat *et al.*, 2013; Field *et al.*, 2012; Masson-Delmotte *et al.*, 2021; Fowler *et al.*, 2021; Martel, Brissette, Lucas-Picher, Troin & Arsenault, 2021). This projected increase can be attributed to the increase of water holding capacity of the atmosphere: Based on Clausius–Clapeyron rate, the water holding capacity of atmosphere increases by 7% per 1° increase of temperature (Molnar, Fatichi, Gaál, Szolgay & Burlando, 2015; Westra *et al.*, 2014). This however cannot directly be translated into precipitation, as the amount of available humidity required for precipitation complicates the relationship (Lochbihler, Lenderink & Siebesma, 2017; Yin *et al.*, 2018). Depending on moisture availability, warming can cause intensification of convective storms with daily or sub-daily scales (Westra *et al.*, 2014).

Considering that precipitation is an essential driver of flood events, different reactions from small and large-scale catchments should be expected: for small catchments, the response time is short and the maximum peak flow can be deduced from a storm with a duration equal to the longest flow path in the catchment (Blöschl, 2022a). Given that the short period of convective rainfall matches the residence time of small catchments, these catchments are the most vulnerable to flooding from convective rainfall, which is expected to increase due to climate change (Viglione & Blöschl, 2009; Viglione *et al.*, 2016; Breinl, Lun, Müller-Thomy & Blöschl, 2021). Regarding large catchments of more than a thousand square kilometers, it is unlikely that

a convective storm leads to a flooding event considering the larger storage capacity and longer travel time (Contractor, Donat & Alexander, 2021). For Nordic snow-dominated catchments, since global warming will likely reduce the amount of snow that accumulates, the magnitude of the spring freshet is expected to diminish. However, even for those catchments, it is anticipated that the frequency and magnitude of convection-driven summer-fall floods, to which small catchments are sensitive, will increase (Donat *et al.*, 2016).

High temporal resolution time series (hourly) of historical data to evaluate the trend of convective storms and consequent floods are difficult to find. A common practice is therefore to use a climate modeling chain and perform simulations at high spatio-temporal resolutions (e.g. Swain *et al.*, 2020; Do *et al.*, 2020). Regional Climate Models (RCMs) offer such high-resolution time series at a local scale (Mearns *et al.*, 2017; Leduc *et al.*, 2019). Moreover, the incorporation of convective parameterization has enhanced their capability to capture convective storms Prein *et al.* (2015); Mooney, Broderick, Bruyère, Mulligan & Prein (2017). More recently, large-ensemble RCM datasets have received attention (Martel, Mailhot & Brissette, 2020b; Sanderson, Oleson, Strand, Lehner & O'Neill, 2018; Aalbers, Lenderink, van Meijgaard & van den Hurk, 2018). Large-ensembles are generated by running RCMs several times, each time with slightly different initial conditions (Deser, Knutti, Solomon & Phillips, 2012a; Deser, Phillips, Bourdette & Teng, 2012b). Multiple values are calculated per time-step, which eliminates the need to fit a parametric distribution on the dataset to compute extreme flows (Martel *et al.*, 2020b; Faghih *et al.*, 2022).

Hydrological models are the last component of a hydro-climate modeling chain (Sidle, 2021). Proportional to the growth of computational power, process-based hydrological models are increasingly used for impact studies (Zhang, Nan, Yu, Zhao & Xu, 2018; Dembélé, Hrachowitz, Savenije, Mariéthoz & Schaefli, 2020; Pandey, Khare, Kawasaki & Mishra, 2019; Duethmann, Blöschl & Parajka, 2020; Zhong, He & Chen, 2018). These models solve the governing equations of hydrological processes (with varying degrees of simplification) per grid cell. Distributed models further use routing algorithms to direct accumulated water towards neighboring cells until the basin outlet (Clark *et al.*, 2015, 2017). The advantage of using distributed physics-based hydrological models is to represent the topography, land use, and soil structure in the model,

to obtain a detailed distribution of hydrological variables of the catchment (Refsgaard, 1995). Therefore, these models are useful to study the internal dynamics of state and flux variables (Golden & Knightes, 2011; Gebremicael, Mohamed & Van der Zaag, 2019; Sidle *et al.*, 2017).

Scale issue is the subject of a long ongoing debate in the scientific community (Blöschl & Siva-palan, 1995; Blöschl *et al.*, 2019). Despite numerous types of research to understand runoff generation processes, there are still unknowns about upscaling from profile scale (1m) to catchment scale and beyond. For example, while the infiltration excess is the governing process at profile-scale (Horton, 1933), the spatial connectivity of hydrological processes has a central contribution in runoff generation at the hillslope scale (Dunne & Black, 1970; Noguchi, Tsub-oyama, Sidle & Hosoda, 1999; Sidle, 2006). Moreover, the contribution of overland connectivity in flow generation and sediment transport and their feedback loop add to the non-linearity of runoff generation (Gomi, Sidle & Richardson, 2002; Jencso *et al.*, 2010; López-Vicente *et al.*, 2017; Koci *et al.*, 2020). The non-linearity from hillslope- to catchment scale is also significant, as the traditional bottom-up Freeze & Harlan (1969) approach to linearly combine all hillslopes so as to compute catchment response has been challenged. Dooge (1986), for example, suggests that a catchment is an "organized complex system", in the sense that the development of co-evolutionary surface and subsurface patterns contributes to catchment drainage and runoff generation (Sivapalan & Blöschl, 2015; Savenije & Hrachowitz, 2017). Adapting the bottom-up approach to these criticisms, there were efforts to combine the hillslope's responses by considering the spatio-temporal covariance of hydro-climate variables for flood simulations (Woods & Sivapalan, 1999; Viglione, Chirico, Woods & Blöschl, 2010).

The spatio-temporal discretization of distributed models can potentially modify land-use and soil structures and result in variations of hydraulic conductivity as well as surface and subsurface hydrological connectivity (Beven, 2000). This can potentially lead to variations in the peak flow or seasonal flow. Many studies have explored the effect of land-use change on streamflow (Singh *et al.*, 2015; Li *et al.*, 2019; Yang *et al.*, 2019a; Tavangar *et al.*, 2019). Using more than one land use scenario is a common approach to studying land use change impacts (Breuer *et al.*, 2009; Huisman *et al.*, 2009; Viney *et al.*, 2009; Bormann *et al.*, 2009). The results show that land use

change can increase/decrease the peak flow, depending on catchment size and/or soil structure. Conversely, paired catchment studies have demonstrated that land-use changes can modify mean seasonal streamflow but has minor effects on the peak flow (Brown, Zhang, McMahon, Western & Vertessy, 2005). The effects of spatio-temporal discretization using process-based models have rarely been investigated for natural catchments. Most previous studies concentrated on urban catchments, with a high degree of impermeability and small size (e.g. Cao *et al.*, 2020a; Krebs *et al.*, 2014; Zhou *et al.*, 2017; Cao *et al.*, 2020a). In this context, multiple studies have shown that variation of spatio-temporal resolution can reorient flow direction and significantly change the flow peak Zhou *et al.* (2017); Ichiba *et al.* (2018); Warsta *et al.* (2017).

Markhali, Poulin & Boucher (2022) have shown that the spatio-temporal discretization of a catchment in a model can affect the representation of surface and subsurface hydrological processes in that model and generate a significant variation in the distribution of hydrological variables including streamflow. Such variations are most important in flat catchments or catchments with considerable human intervention (i.e., agricultural lands). The present study focuses on extreme summer-fall flow using a hydro-climate modeling chain. A range of catchments with different surface areas (from below 200 km to more than 1500 km^2) are selected to facilitate the investigation of the combined impacts of climate change and the spatio-temporal discretization in the hydrological model. More specifically, we intend to verify the following hypotheses for the catchments at hand:

1. For small catchments ($< 500km^2$), refining temporal resolution of simulation (from daily to subdaily) increases the relative change (from reference to future) of extreme summer-fall flow. Refining temporal resolution will not significantly affect projected extreme summer-fall flow for large catchments ($> 1000km^2$).

2. The change in the spatio-temporal scale of modeling causes variability in the projection of extreme flow. By increasing the catchment size, the contribution of hydrological model and spatial scale in that variability increases, and that of the time-scale decreases.

The hypotheses will be examined by forcing two process-based distributed models with large-ensemble simulated climate data. To examine the impact of spatio-temporal discretization, the

simulations will be performed at different spatial (100, 250, 500, 1000 m) and temporal scales (3- and 24-hour time-steps). The structure of this research is as follows: Section 4.2 provides a detailed explanation of the study area, available data, bias correction method, hydrological models, and the experimental plan. Section 4.3 presents the results of the experiments, which are discussed in Section 4.4. Section 4.5 provides concluding remarks and a suggestion for future works.

4.2   Method and Data

4.2.1   Study Area

The study area includes four catchments located in southern Quebec, Canada (Figure 4.1). These catchments range from less than 200 km^2 to more than 1500 km^2 and were selected from diverse land use and hydrological regions. This helps evaluate catchment responses under climate change based on their size and other characteristics, such as land-use and topography. Table 4.1 briefly describes catchments' characteristics.
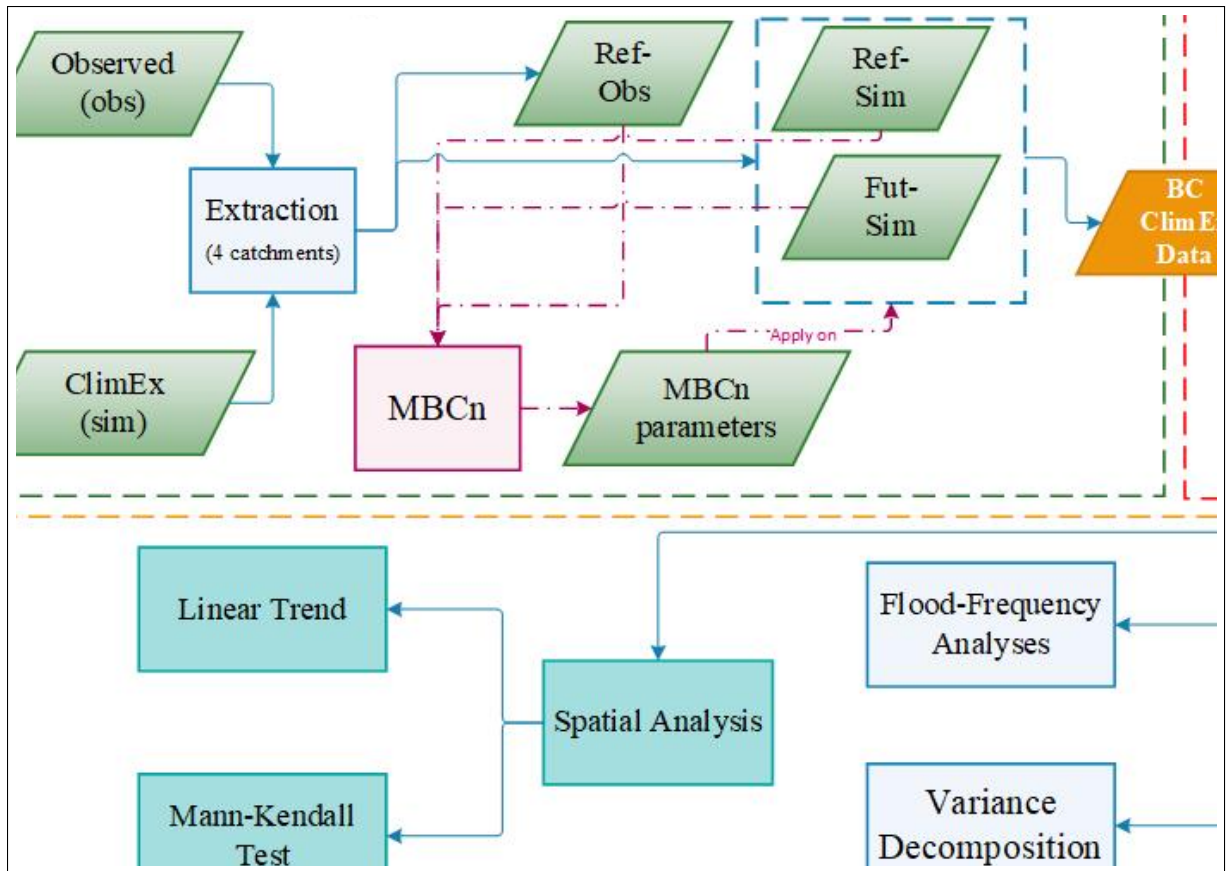
Figure 4.1   Location of the catchments used in this study

Table 4.1   Area and average hydro-climatic characteristics of the catchments used in this study for 2000-2017 period

Number	Name	Area(km^2)	precipitation(mm/yr)	streamflow (m^3/s)	temperature ($°C$)
050135	Croche	1563	1139.36	30.70	2.74
023427	Chaudière	781	1208.65	16.47	3.72
030424	Aux Brochets	584	1329.34	10.52	6.23
023004	Boyer	191	1396.76	4.45	4.15

4.2.2   Datasets

24- and 3-hour observed streamflow series were obtained from the Direction de l'Expertise Hydrique (DEH) of the Ministère de l'Environnement et de la Lutte contre les changements

climatiques (MELCCC) for 2000-2017. Regarding meteorological data, we used the ERA5 (ECMWF ReAnalysis5) gridded dataset to calibrate the hydrological models and simulate streamflow for the present-day climate. (Tarek *et al.*, 2020b) have shown that ERA5 provides an accurate representation of meteorological conditions for catchments located in North America. We also used the ClimEx large ensemble (e.g., Leduc *et al.*, 2019). ClimEx is a 50-member climate dataset, driven by dynamically downscaling the second version of the Canadian Earth System Model large ensemble (CanESM2-LE; Swart *et al.*, 2019), using the 5th generation of the Canadian Regional Climate Model (CRCM5). The simulations are driven by the RCP 8.5 scenario for the period covering 1951-2100, with hourly time steps and a 11° spatial resolution.

4.2.3   Bias Correction

The MBCn (N-dimension multivariate bias correction) (Cannon, 2018) method was selected to bias correct precipitation and temperature time-series extracted from ClimEx. MBCn is an advanced quantile-mapping techniques (Meyer *et al.*, 2019; Cannon *et al.*, 2020). The method transfers all characteristics of the distribution of observations to their simulation according to the climate model. It maintains the trends of projections per quantile, which is essential to accurately assess the impact of climate change (Faghih *et al.*, 2022).

4.2.4   Hydrological models

The following section introduces the hydrological models that are used in this study. Both models are distributed, process-based, and computationally intensive. However, they each have their own methods of representing hydrological processes, and their approach to aggregate the catchment response are also different.

4.2.4.1   WaSiM

WaSiM (Water balance Simulation Model) operates on a raster system (Schulla & Jasper, 2007). The model structure comprises multiple sub-models (e.g., infiltration, evapotranspiration,

snow accumulation and melt, unsaturated zone, etc.) that run on each grid cell and time-step, providing the opportunity to use parallel computing. WaSiM offers two options for calculating the infiltration and percolation: the Topmodel approach or Richard's equation. The first approach is a modified version of the conceptual model Topmodel, following Beven (1997). The second approach is more physically-based and is the one used in this study. All the sub-models that are selected for WaSiM are named in Table 4.2.

4.2.4.2 Hydrotel

Hydrotel is widely used in Quebec for research and operations (e.g., Martel *et al.*, 2020a; Turcotte *et al.*, 2020; Lucas-Picher *et al.*, 2020). In Hydrotel, the catchment is divided into Relatively Homogeneous Hydrological Units (RHHUs) which are hillsopes and include integrated land use class and river segment. Hydrotel is compatible with GIS and remotely-sensed data (Fortin *et al.*, 2001a). A mixture of physical, conceptual, and empirical relationships are used to represent the hydrological processes, which makes Hydrotel slightly less physics-based than WaSiM. For example, the vertical water balance and the representation of soil water content are computed through a sub-routine called BV3C (Bilan Vertical à 3 Couches), which divides the soil column into three layers and controls infiltration, interflow and baseflow, based on a semi-physical moisture accounting equation (Fortin *et al.*, 2001a). Like WaSiM, Hydrotel provides multiple options for representing the hydrological processes of a catchment. Table 4.2 lists the submodels that are used in this study.

Table 4.2   The submodels used to represent the hydrological processes in Hydrotel and WaSiM.

Submodels	Hydrotel	Wasim
Interpolation	Thiessen polygons	Thiessen polygons
Snow melt/accumulation	Degree-Day Method	Degree-Day Method
Potential evapotranspiration	Hydro-Quebec	Hamon (Hamon, 1961)
Real evapotranspiration	BV3C	Richards' Eq.
Infiltration and percolation	BV3C	Richards' Eq.
Channel routing	Kinematic Wave Eq.	Kinematic Wave Eq.

4.2.5   Experimental plan

4.2.5.1   Climate Data Processing

Figure 4.2 shows the experimental plan for this study. The panel on the top left, bounded by the green dashed line, shows the details regarding climate data processing. The first step is the extraction of the simulated and observed meteorological data (temperature and precipitation) for the selected catchments. The reference period for the observed dataset ($Ref - Obs$) spans from 1991 to 2010. ClimEx simulations are also split into reference ($Ref - Sim$) and future ($Fut - Sim$, 2011-2099) periods. In the next step, the 50-member ClimEx (i.e., $Ref - Sim$ and $Fut - Sim$) are pooled together into one long time series per period. This pooling helps to maintain the internal variability of the simulated climate data after bias correction. This is because individual bias correction of each member eliminates the spread of simulations and creates rather similar time series. While addressing internal variability is not among the objectives of this research, maintaining that helps accurate calculation of extreme flows (Faghih *et al.*, 2022). The $Ref - Obs$ and $Ref - Sim$ datasets, which include precipitation and temperature for both the reference and future periods, are further received by MBCn to obtain correction factors based on multi-variate quantile mapping. A single set of correction

factors was computed per calendar month and applied to the simulated climate data. The pooled bias-corrected datasets are reversed back to the 50-member time series, ready to use as the inputs of the hydrological models.
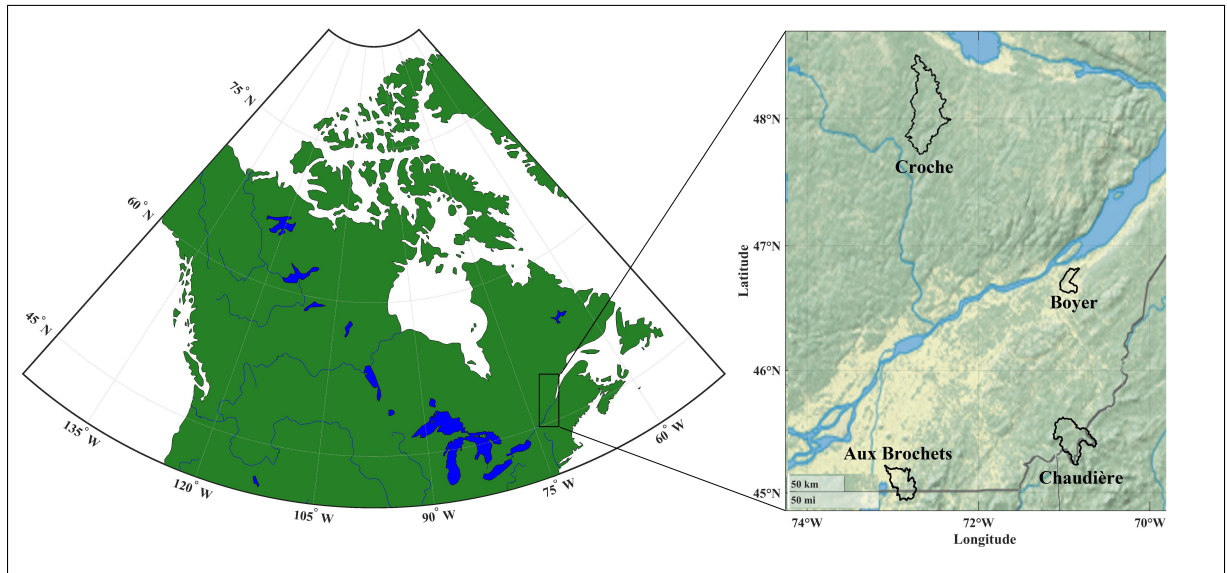
Figure 4.2   Schematic explanation of the experimental plan and methods

4.2.5.2   **Hydrological Simulation**

The hydrological models are calibrated with four spatial (100, 250, 500, 1000 m) and two temporal resolutions (3- and 24-hour). The datasets are split into calibration and validation periods with equal duration. The Dynamically Dimensioned Search (DDS; Tolson & Shoemaker, 2007) with a 0.2 perturbation factor was used to calibrate the models. The DDS technique scales the parameters search space according to a budget specified by the user. Given that both WaSiM and Hydrotel are computationally intensive, this is an advantage over other search methods. In addition, the efficiency of DDS with global parameter perturbations at the beginning and narrowing down the search space by the end of the process has been confirmed in the literature (e.g., Huot *et al.*, 2019).

Based on the existing literature and following experts' recommendations as well as the team who develops and maintains WaSiM, 12 parameters were calibrated, including seven parameters

that are involved with the unsaturated zone subroutine, two parameters linked to potential evapotranspiration, one parameter for snow accumulation and melt, and two parameters for spatial interpolation. The remaining parameters were left to their default values following the WaSiM documentation (Schulla & Jasper, 2007). Regarding Hydrotel, of 28 model parameters, 11 were calibrated, and the others were left to their default values according to Hydrotel's user manual. Out of the 11 calibration parameters, three belong to vertical water balance, six to the snow accumulation and melt routine, and one to the infiltration and interpolation components (see Huot *et al.* (2019) for more details about the parameters).

The Kling-Gupta Efficiency criterion (KGE; Gupta *et al.*, 2009) is the objective function for the calibration of both WaSim and Hydrotel. Compared to other performance criteria, such as the Nash–Sutcliffe efficiency (NSE), the KGE is a better choice as it gives more realistic results for snow-dominated catchment efficiencies. This is because the observed mean is the baseline model for NSE, and for the catchments with high seasonal variability, the measure tends to overestimate modeling skill (e.g., snow melt streamflow; Gupta *et al.*, 2009). Equation 4.1 was used to calculate the KGE

$$KGE = \sqrt{(r-1)^2 + (\frac{\sigma_{sim}}{\sigma_{obs}} - 1)^2 + (\frac{\mu_{sim}}{\mu_{obs}} - 1)^2} \qquad (4.1)$$

where r is the linear correlation between observed and simulated streamflow values, σ_{sim} is the standard deviation of the observations, σ_{obs} is the standard deviation of the simulation, μ_{sim} is the simulation mean, and μ_{obs} is the observation mean.

After obtaining the parameters corresponding with the four spatial resolutions and two temporal resolutions mentioned above, climate simulations from ClimEX were used as inputs to the hydrological models for 1991 to 2099; as mentioned in the top right panel bounded with the red dashed border in Figure 4.2.

4.2.5.3 Analyses

The panel at the bottom of Figure 4.2, with a dashed orange border, shows the analyses and experiments that were carried out to verify the two hypotheses which are the object of this research. To verify the first hypothesis, extreme summer-fall flows are calculated for different spatio-temporal simulations. The streamflow series were split into historical (1991-2010) and far-future periods (2081-2100) to estimate the change of extreme flow under climate change. A 50-member ensemble of simulated streamflows obtained from forcing the hydrological models with ClimEX was pooled together to create a time series comprising 1000 years of data (20 years $\times$ 50 members). This very large ensemble was created to estimate projected yearly extreme flows without the need to fit a parametric distribution. The annual maximum summer-fall flows (July-November) is extracted from the data and an empirical cumulative distribution function is created for both periods (present-day and far future). This allows us to compare the distributions of projected extreme flows in the historical and far-future periods for different combinations of spatio-temporal discretizations (we have four spatial and two temporal resolution that amounts to 8 different combinations). The studied extreme flow values are based on the following percentiles : 50, 90, 95 and 99 (representing 2-, 10-, 20- and 100-year return periods). The procedure regarding pooling and extracting the extreme values is the same as in Martel *et al.* (2020a).

In order to verify the second hypothesis, we use variance decomposition Montgomery (2017) to find the contribution of different factors in the total variance of the projected extremes. Variance decomposition is a simple yet robust and widely applied method (e.g. Addor *et al.*, 2014; Meresa & Romanowicz, 2017; Wang, Huang, Fan & Li, 2020; Meresa *et al.*, 2022). Equation 4.2 shows the application of the method in this study,

$$\Delta U_{i,j,k} = H_i + S_j + T_k + H_i * S_j + H_i * T_k + S_i * T_k + \epsilon \qquad (4.2)$$

where ΔU is the total variance of projected extreme flow, H_i, S_j, and T_k are different choices of hydrological model, spatial resolution, and time-step, and ϵ represents a Gaussian white noise.

To quantify the change in the streamflow when the spatial resolution varies, annual maximum summer-fall flows were extracted per grid and the linear trends corresponding to those grids were computed for the entire 1991-2100 period. The linear trend analysis has frequently been used for quantifying the change in the climate variables (Barnes & Barnes, 2015; Zhuan *et al.*, 2018; Ding & Steig, 2013). Note that the non-linear quadratic and cubic polynomials produced poor results for this case study. The widely used non-parametric Mann-Kendall trend test (Ali, Kuriqi, Abubaker & Kisi, 2019) was also applied to identify the trend at a 0.05 significance level. In this test, the null hypothesis (H_0) assumes no trend and the alternative hypothesis (H_1) assumes the existence of a trend at a 0.05 significance level.

4.3 Results

4.3.1 Annual Hydrographs

Figures 4.3 and 4.4 show the annual simulated hydrographs for the reference and future periods at 3- and 24-hour time-steps using WaSiM and Hydrotel for the Boyer and Croche catchments. These catchments are the smallest and largest, respectively. In these figures, the ensemble of streamflow simulations is based on the ClimEx dataset, for the 1990-2100 period with various spatial resolutions for both hydrological models ($100, 250, 500$, and $1000m$). The median of each ensemble is displayed as a solid line and the observed data is the dashed black line.

Figure 4.3, a) to d) show the WaSiM simulations with 3- and 24-hour time-steps for the Boyer catchment. The observed data is located inside the spread of simulations and the timing of the peaks is approximately the same for both the simulation and the observations (panels a and c). However, the simulation underestimates the magnitude of the median peak flow. We also want to assess how changing the spatial resolution would affect the simulation of low and high flows. According to Figure 4.3, a) to d), the simulation of low flows is more sensitive to variations in spatial scale than that of high flow. Moreover, this sensitivity also increases by refining the time-step from 24- to 3-hour. Figure 4.3 e) to f) shows Hydrotel simulations. As for WaSim, the observed value is located inside the ensemble's spread (panels e and g). Moreover, the

ensemble's median is closer to the observation than that of WaSiM simulations. With Hydrotel, the simulation of high flows is more sensitive than the simulation of low flows to changes in spatial resolution, which is the opposite behavior of WaSim. Again, this sensitivity to the change of spatial resolution is higher for the 3-hour time-step than for the 24-hour time-step. Comparing future (panels a, c, e, g) and reference (panels b, d, f, h) periods, a forward shift of the spring freshet from mid-April to mid-May with significantly lower amplitude can be seen, regardless of the time-step and spatial resolution, for both Hydrotel and WaSim. Overall, WaSiM shows higher sensitivity to changes in spatial resolution than Hydrotel, which is expected as the model is fully distributed and more physically representative in terms of the vertical water budget in the soil.

Figure 4.4 shows the result of the same exercise, but for the Croche catchment. Panels a) to d) show WaSiM simulations with 3- and 24-hour time steps. Compared to Figure 4.3 for the Boyer catchment, WaSiM (panels a and c) shows more skill, as the medians of all the simulations follow the observations closely. In general, varying the spatial resolution has only minor effects on these simulations, except for the simulations with a 3-hour time-step. Hydrotel's simulations (panels e to h) show an underestimation of peak streamflow when the ensemble median is compared to the observations. This underestimation is larger for the simulations with a 24-hour time-step than for the 3-hour time-step. In terms of spatial resolution, both WaSim and Hydrotel are more sensitive to changing the spatial resolution when the time-step of the simulations is finer. Comparing the future and reference periods, a significant attenuation in the magnitude of the spring freshet and a backward shift in the timing of the peak can be observed for both models. There is also a considerable increase in streamflow in the fall and winter months (November to March) when comparing the present and future periods.

Overall, it is not clear from Figures 4.3 and 4.4, whether there exists a pattern regarding the interaction between catchment size and the choice of hydrological model and spatio-temporal resolution. However, each of these elements can distinctly alter the catchment responses.
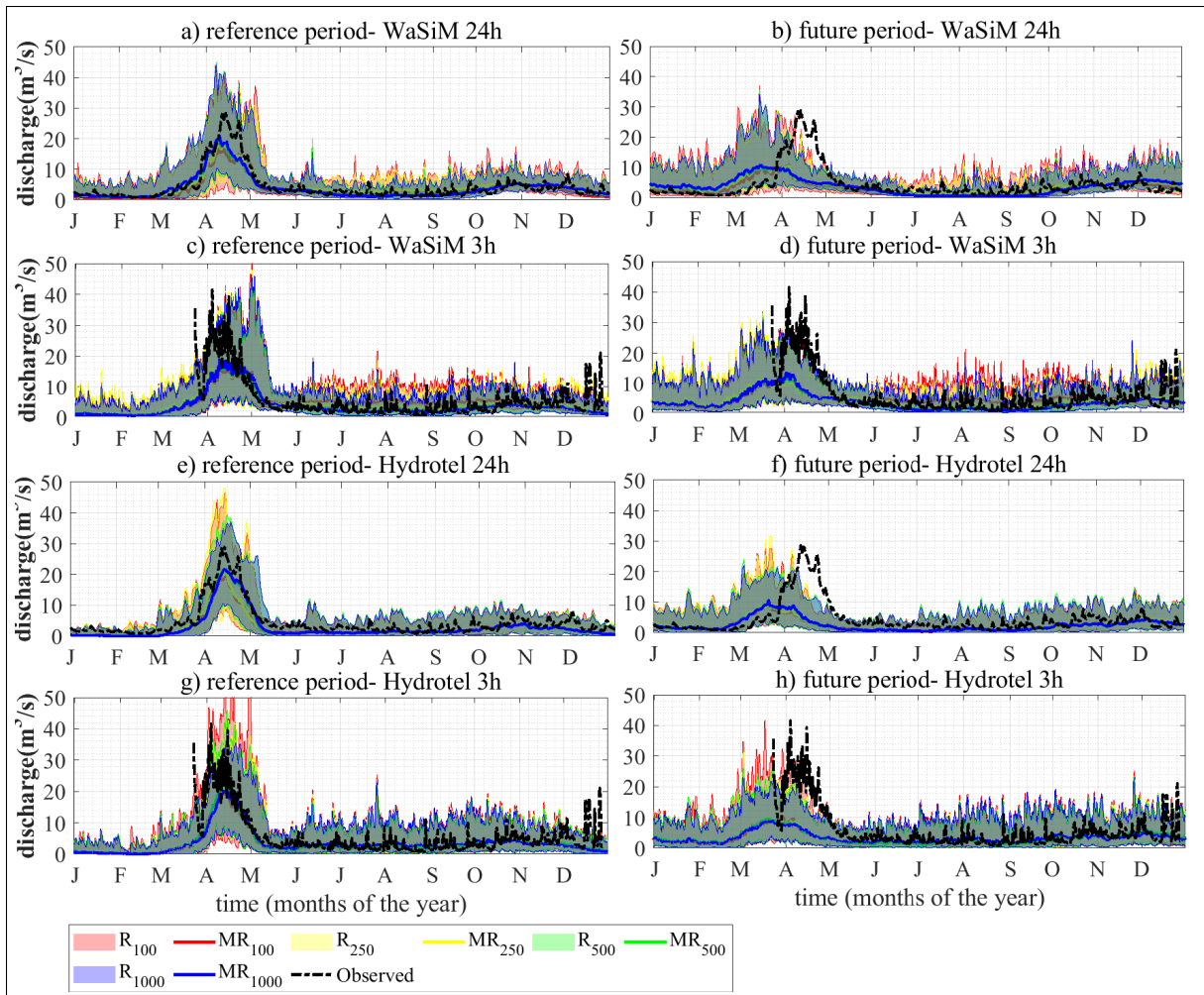
Figure 4.3   Ensemble of annual hydrographs forced by ClimEx dataset per resolution and compared with observed streamflow (dashed black line) for the Boyer catchment. R and the following number represent the spatial resolution in *m*, and MR with the following number represents the median of the ensemble

4.3.2   Spatial distribution of the hydrological variables

Figure 4.5 shows the spatial distribution of actual evapotranspiration (AET) and snow depth (SD) for the Croche catchment. To create this figure, mean annual AET and SD over the reference and future periods were calculated and the relative change between those valuse were obtained. The figure shows that by the end of the century (period 2081-2100), AET will increase by 5 to 10 % according to Hydrotel simulations (panels a to d) and 15 to 30% according to WaSiM
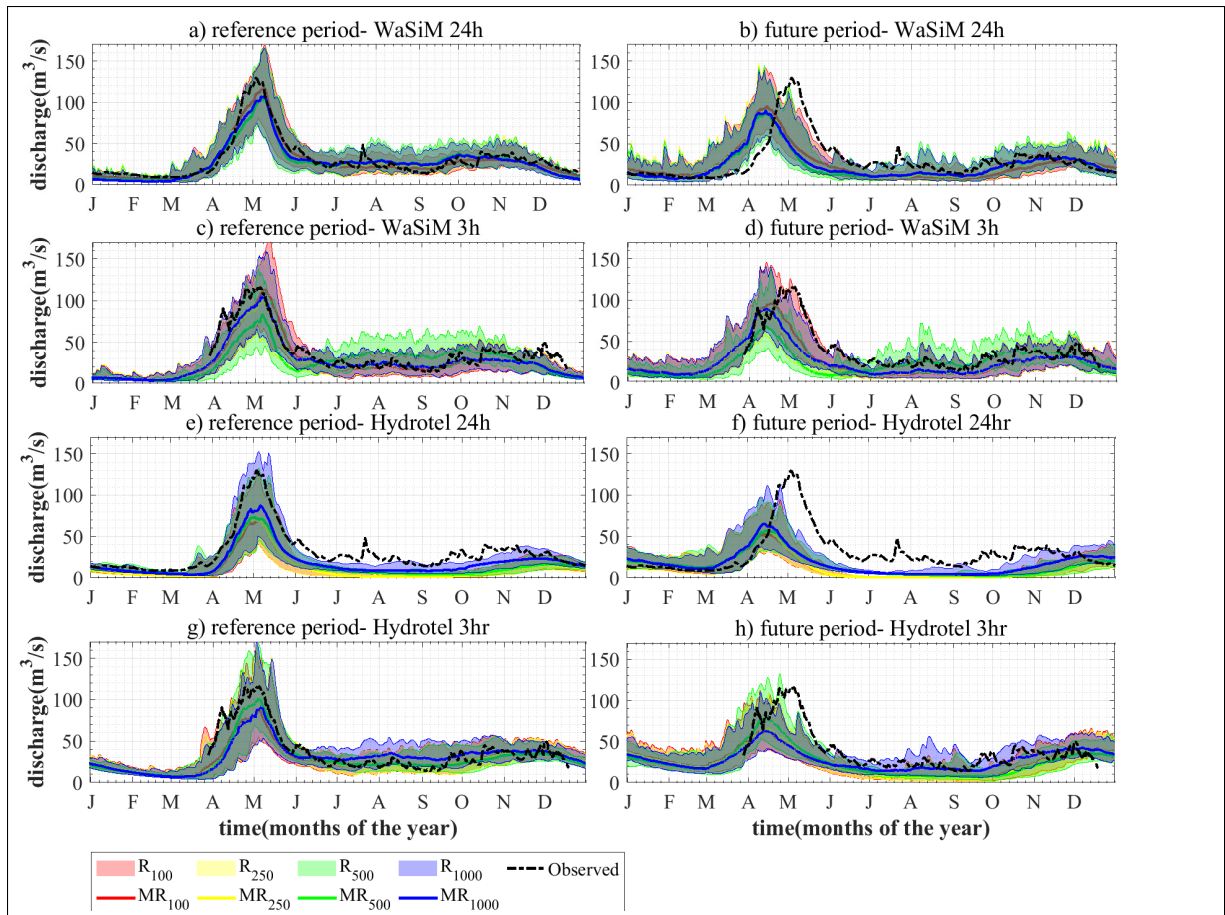
Figure 4.4   Ensemble of annual hydrographs forced by ClimEx dataset per resolution and compared with observed streamflow (dashed black line) for the Croche catchment. R and the following number represent the spatial resolution of simulations in *m*, and MR with the following number represents the median of that ensemble

simulations (panels e to h). A significant negative change in snow depth is observable, as by the end of the century, the average amount of snow on the ground through the winter decreases by around 40% to 50% according to WaSiM (panels m to p) and Hydrotel (panels i to l) simulations. The considerable reduction of spring freshet between 2081-2100, as seen in Figure 4.4, is a result of that reduction in snow depth. Since the amount of snow depth reduction in simulations with Hydrotel is higher than in WaSiM (comparing the third and fourth rows in Figure 4.5), the hydrographs produced by Hydrotlel (Figure 4.4: panels f and h) are more flattened than those produce by WaSiM (Figure 4.4: panels b and d).
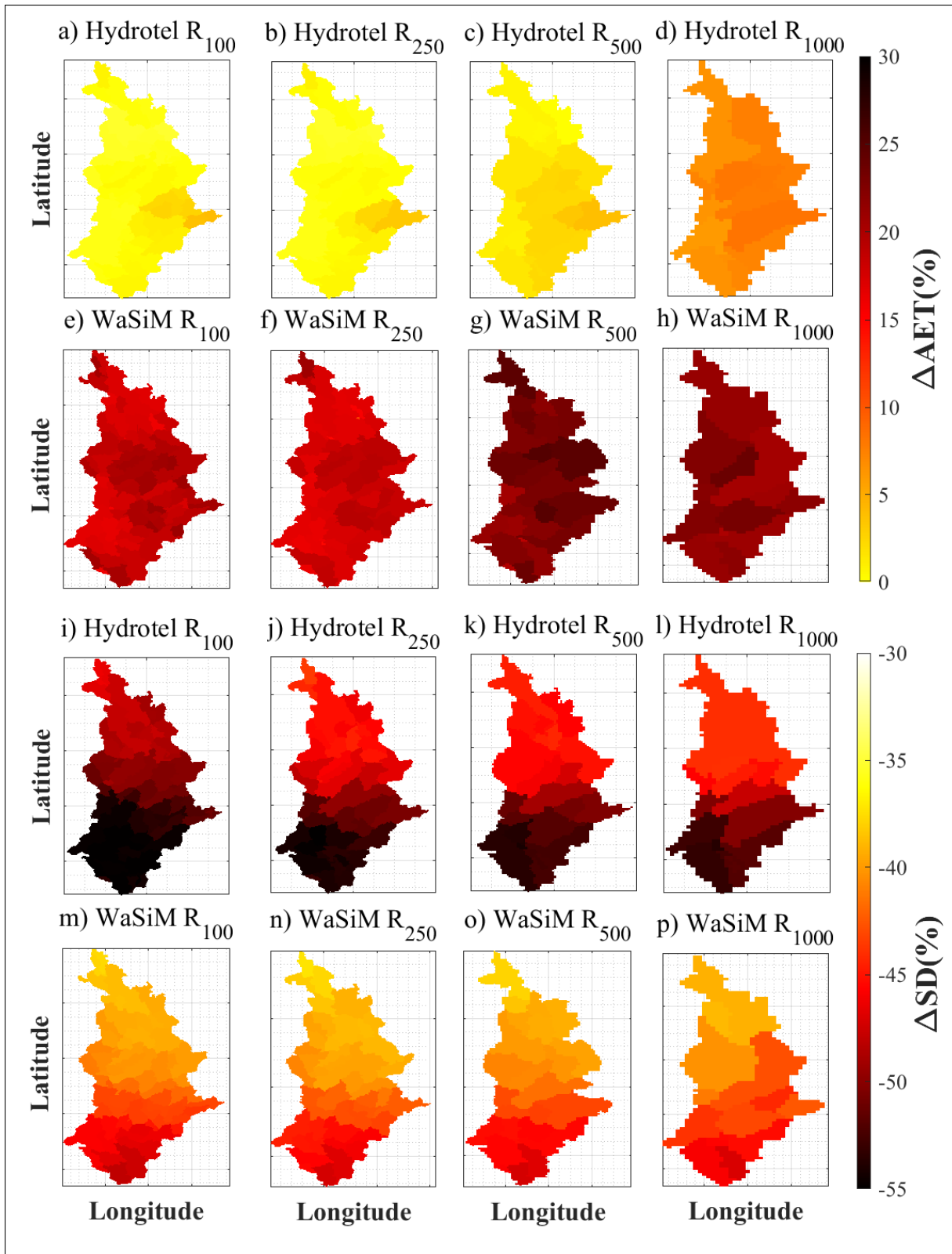
Figure 4.5   Percentage change of actual evapotranspiration (AET) and snow depth (SD) from reference (1991-2010) to far-future (2081-2100) periods for the Croche catchment. R and the following number represent the spatial resolution of simulations in *m*

Changing spatial resolution affects the magnitude of change in the simulation of AET. According to panels a to d and e to f (Figure 4.5), decreasing spatial resolution corresponds with around a 5 to 15% (depending on the hydrological model) increase of change in the AET. For snow depth, changing spatial resolution has no considerable effect on the final results. No significant spatial pattern has been detected for the distribution of AET across the catchment. For snow depth, both models agree on projecting lower values for the southern part with lower altitude illustrating that low-altitude regions are more sensitive to the effect of climate change than high-altitude regions.

4.3.3   Summer-fall extreme flow

Figure 4.6, shows the empirical cumulative distribution function of maximum summer-fall flow simulated by WaSiM and Hydrotel for the reference and future periods per catchment and resolution. The catchments are ordered in terms of size: the first row shows the results for the smallest catchment and the last row shows the results for the largest. Each spatial resolution is identified by a different color and the future and reference periods are shown in dashed and solid lines respectively. The letters W and H represent WaSiM and Hydrotel respectively, with subscript numbers that indicate the time-step of the simulation (e.g., W_{24} is the WaSiM simulation with a 24-hour time-step.)

A pattern regarding the effect of catchment size and the choice of temporal resolution on the change of extreme flow between the reference and future periods is observable. For small catchments (Boyer and Aux Brochets), by refining the time-step of simulation, there are flow quantiles from which the future extreme flow exceeds that of the reference. This is clearer for WaSiM simulations in panels c) and g). For example, in panel c), when the spatial resolution is 100 m, future flows larger than a flood with a 2-year return period (i.e. non-exceedance probability of 0.5) is larger than that of the reference. A similar pattern is also observable for Hydrotel simulations of those catchments (i.e. d and h), even though not as clear as WaSiM simulations. By increasing the catchment size (Chaudière and Croche), the graphs show decreasing magnitude of extreme summer-fall flow between the reference and future periods with the same return periods, regardless of temporal resolution. These observations are in
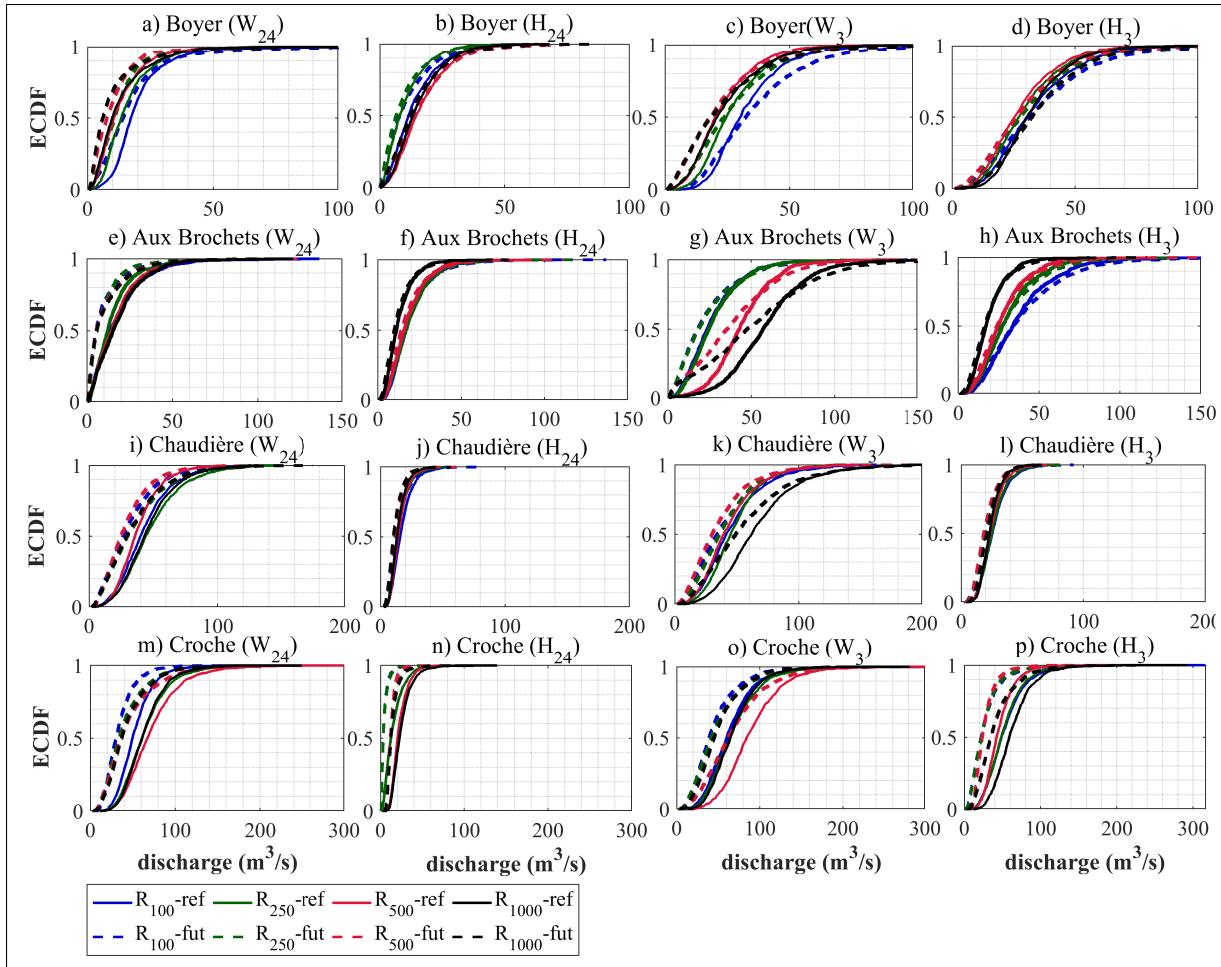
Figure 4.6   Empirical cumulative distribution function (ECDF) of extreme summer-fall flow for reference (ref-solid lines) and future (fut-dashed lines). R and the following number represent the spatial resolution in *m*. W and H are simulations with WaSiM and Hydrotel, respectively and their following numbers represent the temporal resolution in hour

accordance with the first hypothesis of this research. Note that the case of Boyer catchments is complicated for very large return periods (>100-year), as even when the time-step of simulations is 24 hours, the magnitude of a future flood with the same return period exceeds that of the reference period (panels a and b).

For smaller catchments, by changing spatial resolution, the intersection point of future and reference graphs significantly varies. For example, in panel g), with 1000 *m* spatial resolution, the intersection point is equivalent to a flood with the magnitude of a 3-year return period, but

when the spatial resolution is 100 *m*, the intersection point is equivalent to a flood with the magnitude of the 10-year return period. The difference between simulations caused by the change of spatial resolution can also be seen in panels c, k, and o even though the differences between those graphs are smaller. In all cases, whether the time-step of simulation is 3- or 24-hour, simulations by WaSiM have a higher sensitivity to the choice of spatial resolution compared to Hydrotel. These differences illustrate the importance of the choice of spatial resolution and hydrological model. However, no pattern regarding the catchment size and those choices is observable (therefore the second hypothesis cannot be validated from these results).

It is worth mentioning that the counter intuitive appearance of near zero values corresponding with extreme flows in this plot can be attributed to the members that significantly underestimate streamflow simulation.

To further investigate the observations made regarding Figure 4.6, the relative change of extreme flow for specific quantiles (i.e. flood with 2, 10, 20, and 100-year return periods) is driven and presented in Figure 4.7. The results are again ordered according to catchment size. As can be observed, the relative change increases when catchment size decreases. Comparing panels b) and d) with panels a) and c) shows that the magnitude of such increase is higher for the 3-hour time-step than for the 24-hour time-step (in accordance with Hypothesis 1). Moreover, for simulations with a 3-hour time-step, the number of pixels with a positive ratio is higher than with a 24-hour time-step. This demonstrates that the simulated magnitude of flood events in the future increase if a fine temporal resolution is used.

There is no clear pattern regarding the role of spatial resolution in determining the magnitude and direction of change. However, the choice of resolution is not trivial as it can change the magnitude or even direction of the relative change. The choice of model has also important implications as the patterns for Hydrotel and WaSiM, particularly for 24-hour simulations, are different. Comparing panels a) and c), Hydrotel produces more simulations with positive relative change than WaSiM. Again these observations confirm the importance of the choice of spatial resolution and hydrological models, but cannot validate the second hypotheses.
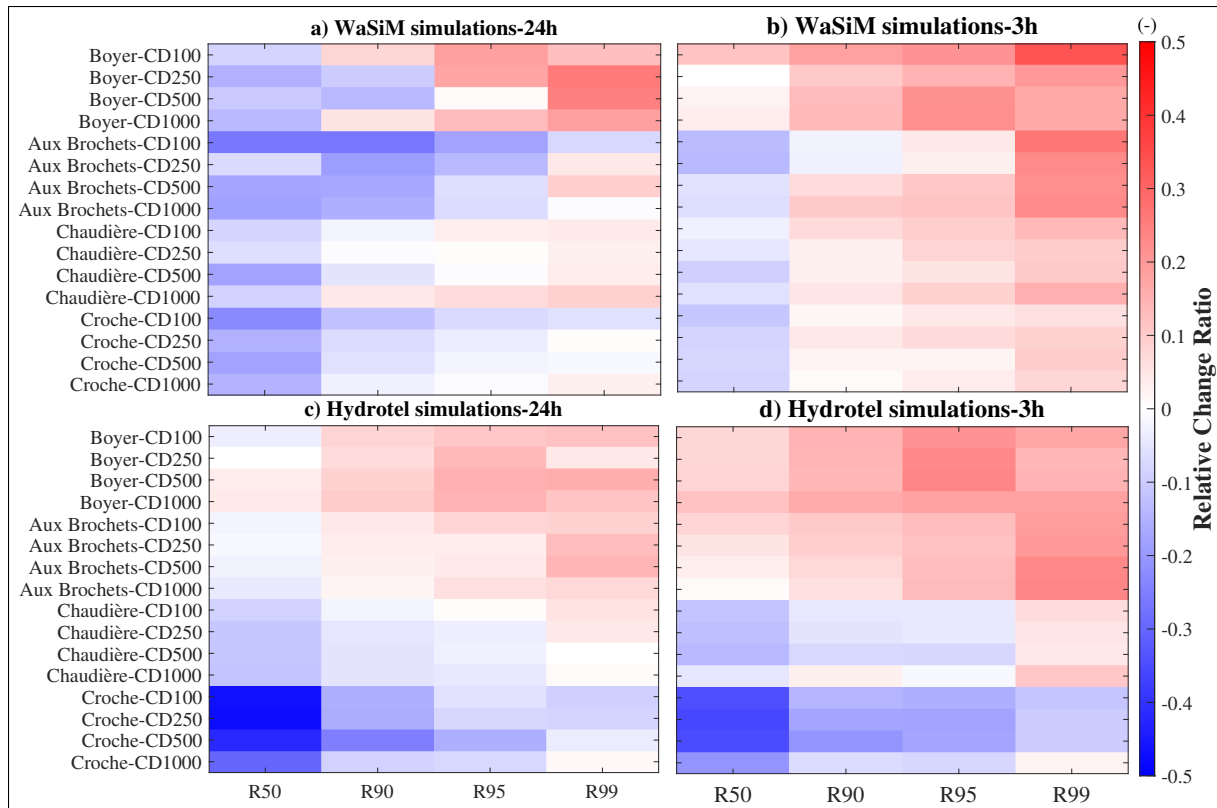
Figure 4.7   Relative change of summer-fall extreme flows (QT50,QT95,...) ordered according to catchment size and spatial resolution

4.3.4   Spatial trend

Figures 4.8 (Boyer) and 4.9 (Croche) show the spatial distribution of the trend for the maximum annual summer-fall flow ,which presented as the percentage of mean annual summer-fall streamflow. For spatial distribution, the simulated streamflow per WaSiM subwatershed is extracted for the entire simulation period (1991-2100). In this figure, R represents the spatial resolution in m, which is followed by the simulation time-step. The hatched area covers the location where the trend is statistically significant at a 5% level ($p - value < 5\%$).

Figure 4.8 shows that a positive trend holds for all simulations with a 3-hour time-step (a to d), regardless of the spatial scale. Moreover, except for $R_{250}(3h)$ (panel a), the trend is significant across most of the catchments. A negative trend emerges across the catchment when the time-step increases (panels e to h), except for the highest (finer) spatial resolution (i.e. e:
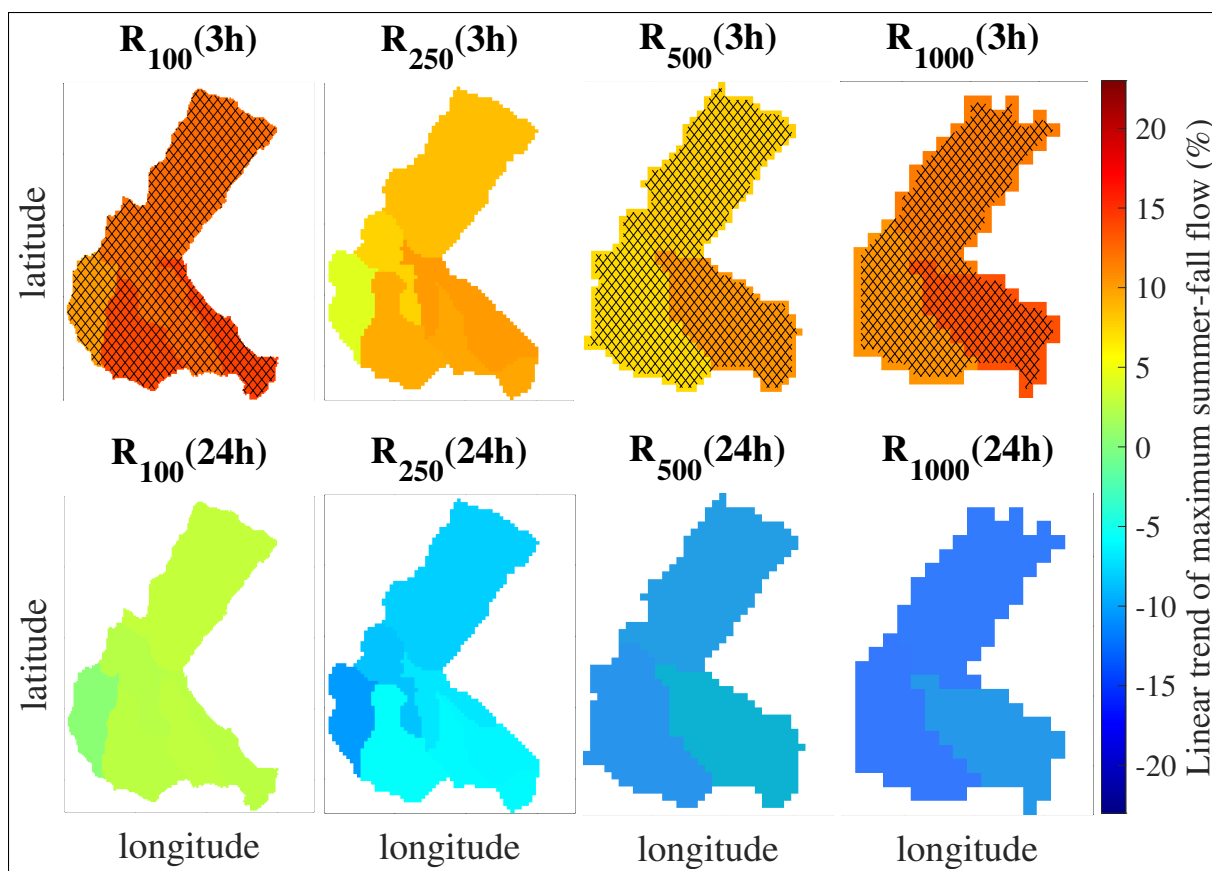
Figure 4.8   Linear trend of the maximum summer-fall streamflow (median member of ClimEx) for the Boyer catchment (all simulated by WaSiM) in terms of percentage of mean summer-fall flow. R and the following number represent spatial resolution in *m*. The hatched area covers the area for which the trend is significant at the 5% level according to the Mann-Kendall test

$R_{250}(24h)$) (hypothesis 1). Changing the spatial resolution has important implications here: the average magnitude of the trend across the catchment varies from larger than +5 to more than +20 % for simulations with 3-hour time-step (panels a to d), and from around +3 to less than -10% of that for daily simulations (panels e to h), illustrating large uncertainties in the projection of high flow (hypothesis 2). There is no distinguishable pattern regarding the relationship between the magnitude and direction of the trend and the spatial resolution.

For the Croche catchment, the spatial distribution of the linear trend is negative regardless of the time step and the spatial resolution of the simulations (hypothesis 1). Similar to the Boyer
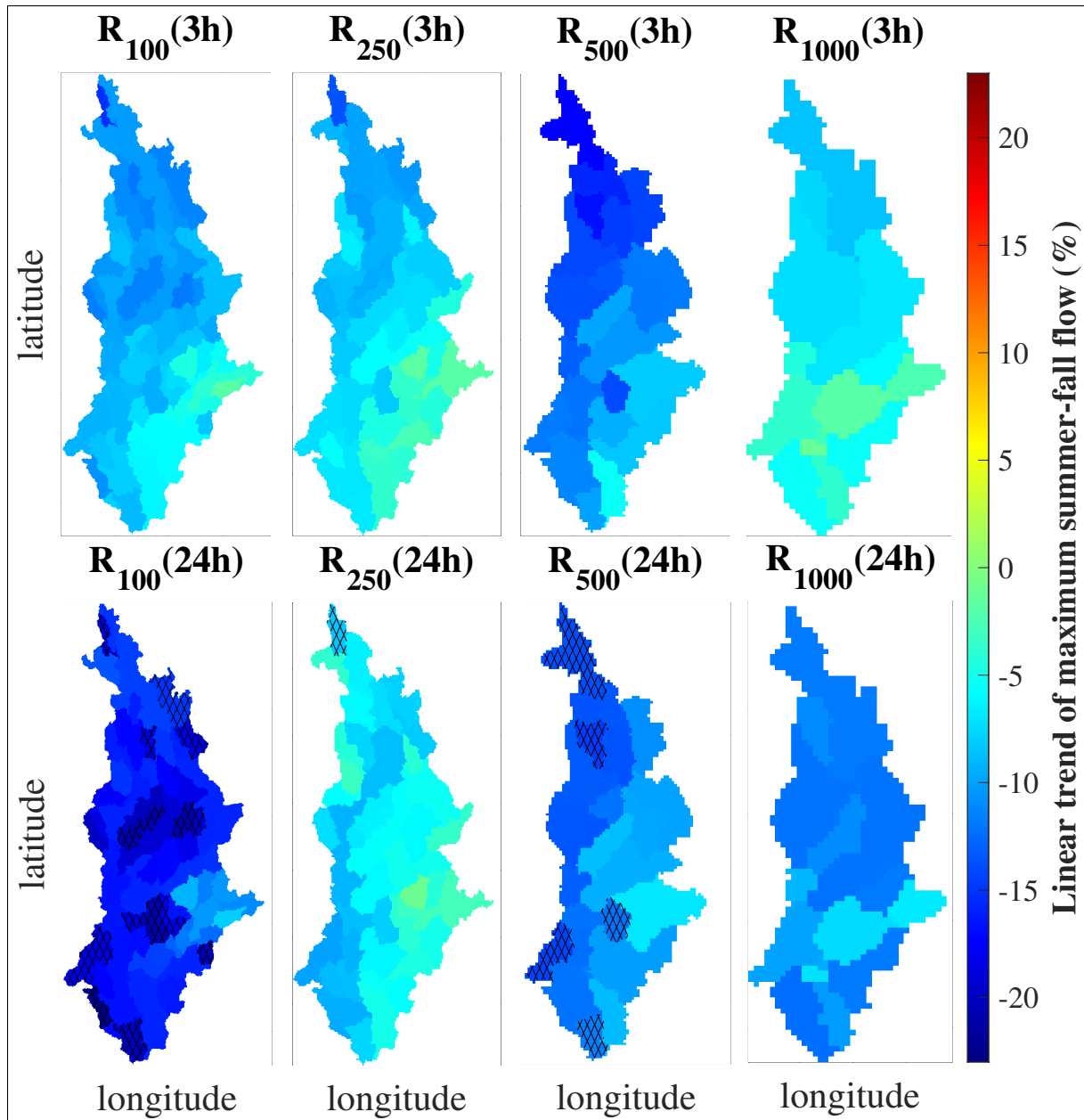
Figure 4.9   Linear trend of the maximum summer-fall streamflow (median member of ClimEx) for the Croche catchment (all simulated by WaSiM) in terms of percentage of mean summer-fall flow. R and the following number represent spatial resolution in m. The hatched area covers the area for which the trend is significant at the 5% level according to the Mann-Kendall test

catchment, changing the spatial resolution of the simulations causes large uncertainties: the magnitude of the trend varies from around -5 to -10 % of summer-fall streamflow for a 3-hour

time-step (panels a to d), and from less than -5 to around -20 % for daily simulations (panels e to h). Note that like the Boyer catchment, no pattern regarding a relationship between the spatial resolution of simulation and the magnitude of the trend is distinguishable. It appears that by changing the time-step from 3- to 24-hour, larger negative trend values (in terms of magnitude) emerge, showing that the subdaily simulations even influence the trend for the large catchment. The influence, however, is not large enough to change the direction of the trend (hypothesis 1). Note that, unlike the Boyer catchment, the trends calculated for the 3-hour time-step are not significant here.

Observations made in this section are in line with the first hypothesis, as refining the time-step of simulation has mostly influenced the small catchment (i.e., Boyer: Figure 4.8) rather than the large catchment (i.e., Croche: Figure 4.9). The second hypothesis cannot be confirmed or rejected with the information provided here.

4.3.5 Analyse of Variance (ANOVA)

Figure 4.10 shows the variance decomposition of the relative change in the extreme summer-fall flow into the contribution of spatial resolution, time-step, hydrological model, and their combinations. Results for smaller catchments are shown on the top side and larger catchments are on the bottom side. The spatial resolution has only a minor contribution to the changes for the Boyer catchment. However, this contribution becomes significant when changing the resolution is combined with other factors (19% of the variance results from changing the spatial resolution and the hydrological model). For the Aux Brochets catchment, the spatial resolution has a larger contribution to the total variance (15%). This is in line with the results from Figure 4.6, where the change of spatial resolution created a large difference between simulations. Interestingly, by increasing the catchment size from 584 km^2 (Aux Brochets) to 781 km^2 (Chaudière) and 1563 km^2 (Croche), the contribution of spatial scale in variability, first significantly drops ($< 1\%$) and then increases back to 14%. This clearly suggests a lack of a clear pattern between catchment size and spatial scale (hypothesis 2 regarding spatial scale cannot be verified). The variance obtained from changing the time step is important for all catchments. But similar to spatial

scale, a clear relationship between catchment size and time-step cannot be found in this context (hypothesis 2 regarding temporal scale cannot be verified). Changing the hydrological model impacts the variance for the largest catchment (Croche) the most, and loses its contribution by decreasing catchment size (Hypothesis 2 regarding the hydrological model can be verified). Note that the combined effect of simultaneously changing the hydrological model and the temporal or spatial resolution can be an important source of variability, but the combined effect of spatial and temporal scale is not as important.
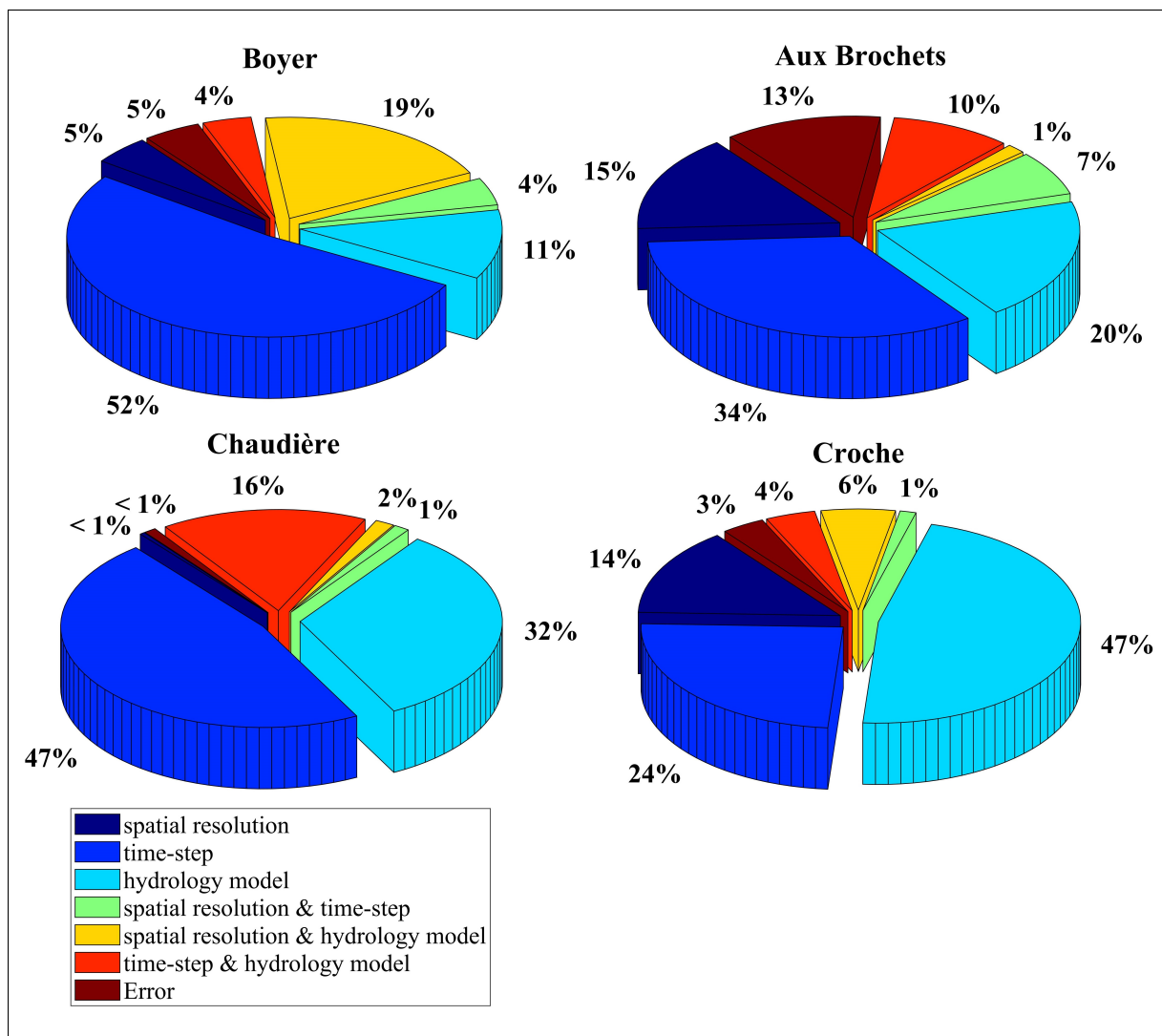
Figure 4.10   Decomposition of the variance for the relative change in summer-fall extreme flows (average of 2, 10, 20, and 100 yr return periods) obtained by ANOVA

4.4   Discussion

This study builds on our previous research (i.e., Markhali *et al.*, 2022) in quantifying the uncertainty linked to the spatio-temporal representation of catchments in hydrological models. In this research, we did not implement the ensemble method by mixing and matching the parameters and catchment descriptors with different resolutions, due to the computational costs of simulating a large-ensemble of long-duration time series. In the previous study, we learned that the uncertainty linked to the catchment heterogeneity is mostly sensitive to the choice of hydrological model, in the sense that the more sophisticated model in terms of representation of hydrological processes (i.e. WaSiM) creates larger uncertainties linked to the catchment heterogeneity compared to less sophisticated model (i.e. Hydrotel).

We focused on quantifying the uncertainty in the projection of extreme summer-fall streamflow. We separated catchments based on their surface area. This was necessary because the flood generation mechanism for small and large catchments are different (Blöschl, 2022b). Small catchments are more sensitive to the infiltration excess runoff, while large catchments are sensitive to the saturation excess runoff (Blöschl, 2022a). The results showed that in fact there are relations between the surface area and the choice of time-step and hydrological model in the final response of the catchments: First, using a finer time-step in simulations resulting in a statistically significant increase in the projection of summer-fall flood hazard in the future for the small but not for the large catchment (Figures 4.8 and 4.9). Second, by increasing the catchment area, the contribution of the choice of hydrological model in the uncertainty increased (Figure 4.10).

The individual contribution of spatial scale is smaller than the other two factors (it is between 1 to 15 % of the total uncertainty according to Figure 4.10). The question is, whether or not variations of spatial scale should be considered in the simulation for flood projection. To answer this question we investigate the response of the Boyer and Aux Brochets catchments to variation of spatial resolution:

Regarding the Boyer catchment, Figure 4.10 shows that the joint contribution of spatial resolution and hydrological model in the variation of extreme summer-fall flow reaches up to 19%, which is the highest among all catchments. Also, in Figure 4.8 e) for WaSim simulations, when the spatial resolution is 100 m, the trend is zero or positive across the catchment. However, by lowering the resolution (panels g, h, i), the trend becomes negative. According to Markhali *et al.* (2022), increasing the spatial resolution causes a nonlinear decrease in the coefficient of interflow storage in WaSim for this catchment. This means that the saturation level of the soil is significantly higher for simulation with a 100 m resolution compared to other choices of spatial resolutions. Because of the high value of soil moisture for the simulations with a 100 m resolution and increasing convective rainfall in the future, there is a positive trend in the simulation of high flow even if with a daily simulation time-step. By decreasing the resolution, the interflow storage increases, leading to lower antecedent soil moisture and consequently a negative trend for high flow in the 24-hour time step.

The Aux Brochets catchment shows the largest sensitivity to the spatio-temporal resolution in flood projection (Figure 4.6). Coarsening the spatial resolution in WaSiM induces modifications to the slopes of this catchment in the model, which in turn causes a reorientation of surface and subsurface flows. This results in soil saturation in a portion of the catchment leading to the outlet (Markhali *et al.*, 2022). High antecedent soil moisture combined with convective storms results in a rapid response of the catchment for simulations with low spatial but high temporal resolutions. The significantly larger magnitude of flood for the simulations with a 3-hour time-step and a spatial resolution of 500 to 1000 m (the red and black lines in Figure 4.6 *b*) could be attributed to the mechanism explained above. The decomposition of the variance for this catchment in Figure 4.10 confirms that the contribution of the spatial scale individually or together with the other factors explains 23% of the total variance, which is higher than two other larger catchments.

Among all the catchments studied here, the Boyer catchment has the maximum human intervention in terms of deforestation and agriculture (Markhali *et al.*, 2022). Also, Aux Brochets is a flat catchment with uneven areas (hills and valleys). This type of topography is more difficult to

represent in hydrological models. Therefore, the model structure and the degree to which that model reflects the details of topographic and land-use characteristics are important factors to consider. This study suggests accounting for the variation of spatial resolution for flat catchments or catchments with high agricultural lands if a distributed hydrological model with a high level of sophistication in representing hydrological processes should be used.

The intensive computational demand of the two distributed process-based hydrological models used in this research limits the number of catchments that could be included. There is an opportunity to work towards generalizing the conclusions of this research by involving a higher number of catchments, with different sizes and land uses. Moreover, adding more models with various structures seems necessary to gain more in-depth knowledge about the effect of the choice of process-based models in flood projection. Furthermore, the recent advances in increasing the spatial and temporal resolution of RCMs are appealing to further investigate the impact of spatio-temporal resolution in climate impact studies. Recent models with a high spatial resolution (<4km) have shown promise in the simulation of convective-driven rainfall (Lucas-Picher *et al.*, 2021). The problem with using these models is the large storage capacity required for storing data (Gutowski *et al.*, 2020). Also, coupling them with distributed hydrological models adds to the computational costs of the modeling. Further advancements in computational power and data storage are required for the application of these models in impact studies (Schär *et al.*, 2020).

4.5   Conclusion

This study investigated the role of spatio-temporal resolution of simulations, the choice of hydrological model, and the catchment size in determining the change of extreme summer-fall flow in the future under climate change. A large-ensemble regional climate model simulation (ClimEx) was bias corrected by multi-variate bias correction (MBCn) and coupled with two distributed hydrological models (WaSiM and Hydrotel) to simulate streamflow over four catchments with different sizes across Quebec. Simulations were conducted for different spatial (100, 250, 500, 1000 m) and temporal (24- and 3-hour time-steps) resolutions. Multiple

experiments were conducted to reject/validate two main hypotheses: 1) For small catchments, by increasing temporal resolution, the simulated extreme summer-fall flow in the future period becomes larger than that of the reference period. 2) The change in the spatio-temporal scale of modeling causes variability in the projection of extreme flow. By increasing the catchment size, the contribution of the choice of hydrological model and spatial scale in that variability increases, and that of the time-scale decreases.

The experiments show that:

1. A pattern regarding catchment size and temporal resolution exists: simulations with 3-hour time-steps (Figures 4.6, 4.7, 4.8) predict that extreme summer-fall flow will increase in the far-future for small catchments, regardless of model and spatial resolution. Therefore, the first hypothesis is verified for this case study. Moreover, the choice of a simulation time step is a major determinant in the variability of flood projection for small catchments and by increasing catchment size, its influence decreases (Figure 4.10). As a result, part of the second hypothesis concerning the relationship between temporal resolution and small catchments is also verified for this case study.

2. For large catchments, the choice of spatial resolution has a larger contribution in the simulation of extreme summer-fall flood (Figures 4.6 and 4.10). This however does not exceed the contribution of the choice of time-step (Figure 4.10). Moreover, if the time-step is 24-hour, it is likely that the spatial resolution changes the direction of the trend for small catchments (Figures 4.8 and 4.9 and section 4.4). Therefore, part of the second hypothesis concerning the impact of spatial resolution on large catchments cannot be verified here.

3. The choice of a hydrological model can be important for both small and large catchments. It appears that by increasing the catchment's size this choice becomes more important (e.g., Figure 4.8 and 4.10). Therefore, part of the second hypothesis regarding the impact of the choice of a hydrological model on large catchments can be verified here. In all cases, WaSiM shows a higher variance than Hydrotel for streamflow projections (Figure 4.6).

CHAPTER 5

REGIONALIZATION OF A DISTRIBUTED HYDROLOGICAL MODEL USING RANDOM FOREST

Siavash P. Markhali[1] , Marie-Amélie Boucher[2] , Annie Poulin[1]

[1] Department of Construction Engineering, École de technologie supérieure, 1100 Notre-Dame West, Montréal, Québec, Canada H3C 1K3
[2]Civil and Building Engineering Department, Université de Sherbrooke, 2500 Bd de l'Université, Sherbrooke, Québec,Canada J3X 1S1

Abstract

This study uses a machine learning technique (i.e., Random Forest) and a process-based hydrological model for regionalization. The approach shows skill in simulating streamflow in pseudo-ungauged catchments. We ran the model across different spatio-temporal resolutions and investigated three hypotheses: (1) Finer time-step adds more information to the calibrated parameters and improves the efficiency of the regionalization model. (2) The parameters approximated by RF are spatially consistent and transferrable across the spatial scales. (3) Refining the spatial resolution of catchment descriptors (CDs) will improve regionalization skills. The results show that refining the time step significantly improves the modeling skill ($\approx 12\%$ improvement at a significance level of 0.05). In addition, the regionalization model maintains the spatial correlation between CDs and parameters. Finally, for simulations at a daily time step, spatially refined CDs improve the regionalization skill ($\approx 10\%$ improvement).

5.1   Introduction

Process-based distributed hydrological models are suitable tools for understanding the complexity of hydrological processes, which take place on heterogeneous media under ever-changing internal (e.g. land use change) and boundary conditions (e.g. climate change) (Beven, 2011; Blöschl

et al., 2019; Blöschl & Sivapalan, 1995; Fatichi *et al.*, 2016). Corresponding with the evolution of computational power, increasingly sophisticated distributed models have been applied across various hydrological problems including flood forecasting, climate change impact assessment, and analyses of hydrological processes at different spatio-temporal scales (e.g. Addor *et al.*, 2014; Blöschl *et al.*, 2008; Kumar *et al.*, 2013; Rakovec *et al.*, 2016; Thober *et al.*, 2019; Martel *et al.*, 2020a). However, for most process-based hydrological models, some parameters cannot be directly determined and therefore, the generally accepted practice is to calibrate model parameters using observed data (Fatichi *et al.*, 2016). Still, in many locations, even in developed countries, there is a lack of observed streamflow data (at the outlet or at internal locations of the catchments), or it is unreliable due to various difficulties (e.g. inaccessibility of the location, extreme weather conditions, vandalism etc.). (Sivapalan, 2003; Guo, Zhang, Zhang & Wang, 2021)

To stress the practical importance of this issue, the International Association of Hydrological Sciences (IAHS) has declared the period of 2003-2012 to be the decade of prediction in ungauged basins (PUB Sivapalan, 2003). The Model Parameter Estimation Experiment (MOPEX) project (Duan *et al.*, 2006) is another example of an international large-scale initiative on the topic of regionalization. Even though this decade of PUB and the MOPEX project have fostered significant progress (e.g. Razavi & Coulibaly, 2017; Hrachowitz *et al.*, 2013; Parajka *et al.*, 2013), there still exists a need to improve regionalization techniques. The objective of any regionalization technique is to find a relationship between a model parameters and catchment characteristics, which can then be used to estimate the parameters for ungauged catchments. This relationship can further be extrapolated to other modeling elements (i.e. catchments, sub-catchments, hydrological response units).

In general, regionalization techniques can be classified into similarity-based methods (e.g. Vandewiele & Elias, 1995; Randrianasolo, Ramos & Andréassian, 2011; Samuel, Coulibaly & Metcalfe, 2011; Yang, Magnusson & Xu, 2019b; Arsenault, Breton-Dufour, Poulin, Dallaire & Romero-Lopez, 2019), and regression-based methods (e.g. Abdulla & Lettenmaier, 1997b; Wagener & Wheater, 2006; Teutschbein, Grabs, Laudon, Karlsen & Bishop, 2018). The underlying

assumption for similarity-based methods is that model parameters are transferable between catchments with similar physical characteristics or on the basis of proximity; Yet, regression-based methods rely on finding a relationship between catchment descriptions (CDs) and model parameters (Razavi & Coulibaly, 2013, 2017; Guo *et al.*, 2021)

In terms of procedure, Samaniego *et al.* (2010) classified the approaches into post regionalization and simultaneous regionalization techniques. For post-regionalization, the links between the parameters of a hydrological model and CDs are established after calibration over a set of reference (gauged) catchments using multivariable predictor-predictant analyses such as regression. Those links are further cross-validated over the test catchments (e.g. Abdulla & Lettenmaier, 1997a,a; Parajka, Merz & Blöschl, 2005; Wagener & Wheater, 2006; Heuvelmans, Muys & Feyen, 2006). For simultaneous regionalization, an a priori relationship between the parameters and the CDs is assumed and takes the form of a transfer function. Then, the parameters of that function are calibrated and coupled with the hydrological model. In this line of research, Hundecha & Bárdossy (2004) used simultaneous regionalization techniques by defining functional relationships between model parameters and catchment descriptors. Following up studies such as Götzinger & Bárdossy (2007); Hundecha, Ouarda & Bárdossy (2008); Pokhrel, Gupta & Wagener (2008) and Troy *et al.* (2008) applied similar strategies.

Regardless of the technique used for a specific regionalization problem, a loss of modeling efficiency from calibration to regionalization is to be expected. This is because, first, a set of calibration parameters can only be one solution among many for the calibration problem. Therefore, using a specific set of model parameters out of all possible solutions might not necessarily provide complete information for establishing the CDs-parameters relationship required for regionalization (Saadi *et al.*, 2019; Olden & Poff, 2003). Second, the conventional time-space averaged catchment descriptors (e.g., mean temperature or precipitation, or mean elevation over catchments) may contain only limited information, as these predictors do not represent the spatial variability of soil type, land cover, and physiographic characteristics of the catchments (Merz *et al.*, 2020). As a result, a transfer function derived from such sets of

predictors-predictands might not suitably explain the relationship between the parameters and CDs (Kim & Kaluarachchi, 2008).

Another concern regarding regionalization is the a priori relationship assumed between model parameters and CDs. These assumptions are typically based on process understanding, expert knowledge, and empirical evidence (Hrachowitz *et al.*, 2013). For example, Hundecha & Bárdossy (2004) related parameters for snowmelt to land use, and evaporation to soil type and land use. (Samaniego *et al.*, 2010), assumed a more complex relationship between parameters and CDs using pedotransfer functions. Such assumptions however are ambiguous. First, the link between parameters and CDs is not clear in most cases (Merz & Blöschl, 2004). Second, there are different catchment descriptors which are correlated (e.g., precipitation and physiographic data) and provide similar degree of information, resulting in equal power of prediction (Merz *et al.*, 2020). Hence, the choice of CDs that control a specific parameter is not clear, making it difficult to constrain parameters with hydrologically reasonable transfer functions.

Machine Learning (ML) techniques can learn multi-variable predictor-predictand relationships (Shen, Chen & Laloy, 2021; Tyralis, Papacharalampous & Langousis, 2019). Having higher degrees of freedom relative to traditional regionalization methods, ML techniques are capable of detecting non-linear relationships between predictors and predictands through training over large numbers of samples (Nearing *et al.*, 2020). Thus, the need to constrain the parameters with transfer functions, which is laborious and often ambiguous, can be eliminated (Merz *et al.*, 2020). ML has been used in hydrology mainly for prediction and benchmarking (Hsu, Gupta & Sorooshian, 1995; Abramowitz, 2005; Best *et al.*, 2015; Nearing, Mocko, Peters-Lidard, Kumar & Xia, 2016; Kratzert *et al.*, 2019a; Kratzert, Klotz, Hochreiter & Nearing, 2020). Different clustering techniques have also been used for the purpose of catchment classification occasionally, addressing PUB under the context of physical similarity (Papageorgaki & Nalbantis, 2016; Kanishka & Eldho, 2020). The dominant ML algorithm used in hydrology and water resources sciences is the multi-layer perceptron (Shen, 2018).

Despite its applications in the academic community, ML techniques have limited use in the operational community Abrahart *et al.* (2012); Kirchner (2006). This is because of ML models' "black box" nature, in which the internal processes between inputs and outputs remain hidden or complex to track. Transparency and clarity are necessary for decision-making (Boucher, Quilty & Adamowski, 2020; Rudin, 2019). Water managers prefer to have the "right answers" for the "right reasons" to avoid unknown risks (Kirchner, 2006). One approach to close the gap between these two communities is to use ML techniques in parallel with hydrological models. This helps preserve the physical reality and interpretability of the modeling, while "the underlying physical processes" that hydrological models cannot typically capture are approximated by numerical ML techniques (Kasiviswanathan, He, Sudheer & Tay, 2016).

ML techniques are well suited to calculate the hydrological model parameters, for which a tangible physical relationship is not available. Random forest (RF; Breiman, 2001), is a powerful ML technique, which has found relatively limited use in hydrology (Tyralis *et al.*, 2019). Some examples include flood and drought analyses (Anderson, Lucas & Bonfils, 2018; Bachmair, Svensson, Hannaford, Barker & Stahl, 2016; Muñoz, Orellana-Alvear, Willems & Célleri, 2018; Sultana, Sieg, Kellermann, Müller & Kreibich, 2018), analyses of hydrological signature and flow regime (Addor *et al.*, 2018; Snelder *et al.*, 2009; Balázs, Bíró, Dyke, Singh & Szabó, 2018). RF is a supervised learning method based on regression trees (Breiman, 2001). The method belongs to the ensemble learning class, which uses bagging (bootstrap and aggregation), decorrelation and randomization techniques (Tyralis *et al.*, 2019). RF is interpretable to some extent, as the algorithm is based on decision trees, in which the flow of information is trackable. Additionally, RF is fast, stable, and it resists overfitting (Boulesteix, Janitza, Kruppa & König, 2012). The technique can handle small to large sample sizes even if the predictors are highly correlated (Ziegler & König, 2014). Such properties are useful to regionalize hydrological model parameters in the case of computationally intensive hydrological modeling.

The present research aims at the regionalization of a process-based distributed hydrological model. Given that distributed models are sensitive to the heterogeneity of the system (Markhali

et al., 2022), we study the effects of different temporal resolutions and their interaction with spatial discretization in regionalization based on the following hypotheses:

   I.   Finer time-step adds more information to the calibrated parameters and improves the efficiency of the regionalization model.

   II.   The parameters approximated by RF are spatially consistent and transferrable across the spatial scales.

   III.   Refining the spatial resolution of CDs will improve the regionalization skill.

The structure of the paper is as follows: Section 5.2 introduces material and methods including data, study area, and the methodologies used for calibration of hydrological model and building RF model. Section 5.3 provides the results and discussion and Section 5.4 gives a summary and the conclusions regarding our hypotheses.

5.2   Material and Methods

5.2.1   Hydrological model

Hydrotel, a semi -distributed and computationally intensive physics-based model, is used for this research (Fortin *et al.*, 2001a,b). The model is used by the Direction de l'Expertise Hydrique (DEH) as part of their hydrological and flood forecasting system across the province of Quebec. For spatial inputs, the model receives GIS-based gridded data including land cover, soil type, and Digital Elevation Model (DEM) rasters as well as river network and lake polygons. The grids are further aggregated into multiple Relatively Homogeneous Hydrological Units (RHHUs), which are hillslopes . To represent the hydrological processes, the model offers options through various submodels providing a flexibility to the modeling practice. As Figure 5.1 illustrates, Hydrotel consists of six main submodels including interpolation of meteorological data, vertical water budget, snow melt/accumulation, potential evapotranspiration, surface routing, and channel routing. A mixture of empirical, conceptual, and physical relationships constitute the governing equations to represent the processes and sub-processes. Overall, 27 parameters need to be specified. Some of these parameters can be fixed for simulations using expert knowledge

sensitivity analyses (Huot *et al.*, 2019) . However, other parameters must be calibrated. The list of the parameters calibrated in this study along with their descriptions are provided in Table 5.1.
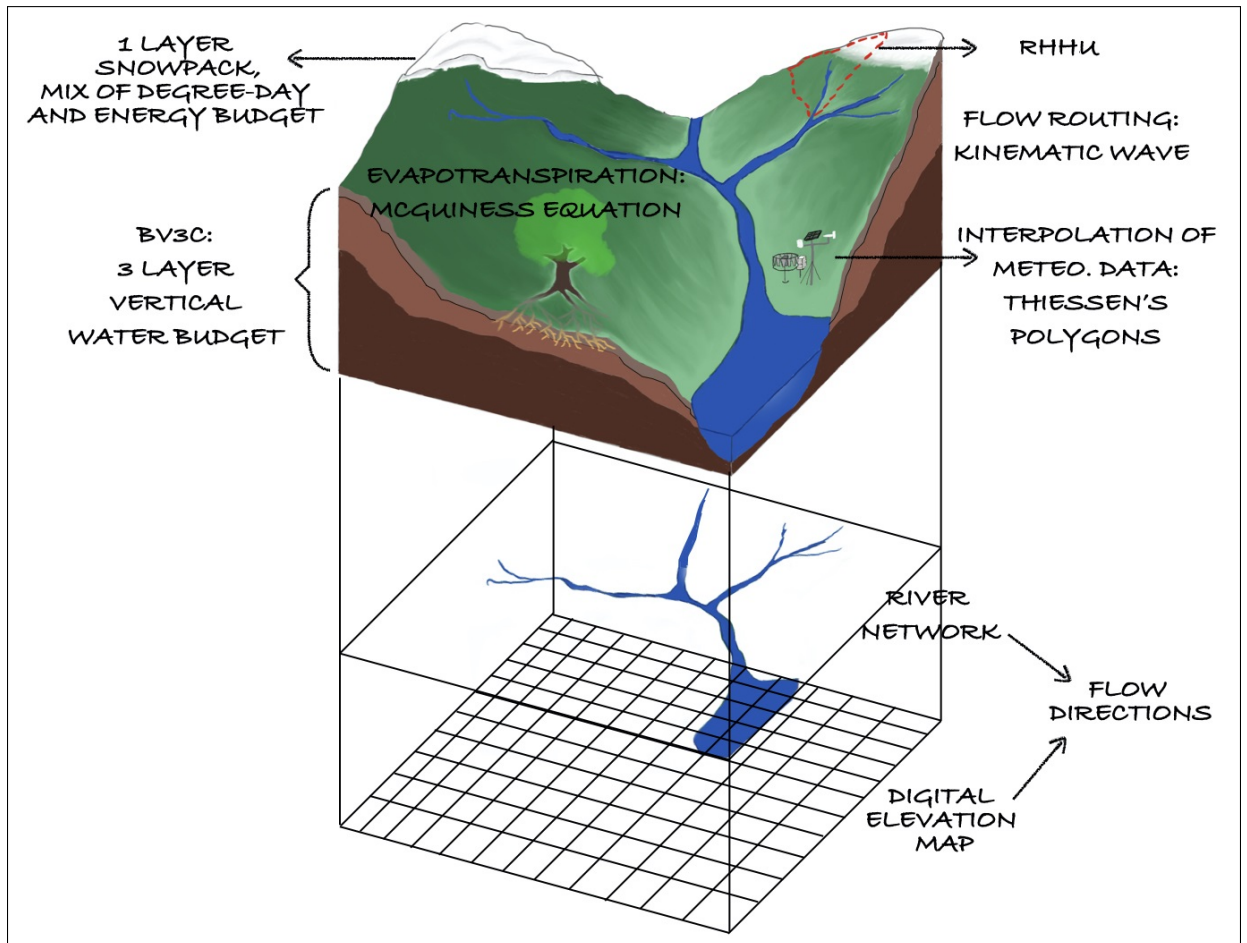
Figure 5.1   Schematic presentation of the hydrological processes and sub-process embedded in Hydrotel

5.2.2   Study Area and Data

We selected 171 catchments from the southern and eastern parts of the province of Quebec in Canada. These catchments are selected among around 400 catchments available in the database provided by the DEH (http://www.cehq.gouv.qc.ca/hydrometrie/historique_donnees/default.asp). The database includes geographical coordinates and some physiographic information of the catchments as well as observed discharge in daily and hourly time-steps. We eliminated

Table 5.1   Hydrotel's calibrated parameters and their description

Parameter	Description (unit)	Submodel
L1	First layer thickness (m)	vertical budget (BV3C)
L2	Second layer thickness (m)	vertical budget (BV3C)
CR	coefficient of recession (m/hour)	vertical budget (BV3C)
MTD	melt threshold deciduous (°C)	snow melt/accumulation(degree-day)
MTN	melt threshold non-forest (°C)	snow melt/accumulation(degree-day)
MTC	melt threshold coniferous (°C)	snow melt/accumulation(degree-day)
MRD	melt rate deciduous (mm/d per °C)	snow melt/accumulation(degree-day)
MRN	melt rate non-forest (mm/d per °C)	snow melt/accumulation(degree-day)
MRC	melt rate coniferous (mm/d per °C)	snow melt/accumulation(degree-day)
CET	coefficient of optimization (-)	Potential evapotranspiration(Mcguiness)
TSR	transition from snow to rain (°C)	Interpolation (Thiessen polygons)

catchments for which all hydrometric stations have been closed prior to 1990, or those with more than 40% of missing streamflow data in the calibration periods. Figure 5.2 demonstrates the spatial distribution of the 171 catchments across the province of Quebec.

Observed gridded meteorological dataset (GC3h dataset- precipitation, and maximum and minimum temperature from 1990-2018) produced by Ministère de l'Environnement et de la Lutte aux Changements Climatiques by kriging the measurements from over 350 ground stations on a 10km by 10km grid is used to produce the dataset over fixed grids. The fine spatio-temporal resolution (3-hour temporal resolution and 10 km spatial resolution) of the dataset is an advantage over similar products, as the resolution of the modeling is of the interests of the present study.

5.2.3   Experimental Setup

Figure 5.3 schematically demonstrates the methodology developed to test the hypotheses laid out in the Introduction section. First, the hydrological model is calibrated over the selected catchments for both 24- and 3- hours time-steps to obtain calibrated parameters. Afterwards, the RF regionalization model is built using the calibrated parameters and CDs as training data according to the first column of Figure 5.3: Building Random Forest (RF) Model. To test hypothesis one, the outputs of the RF-based regionalization model are further used as inputs
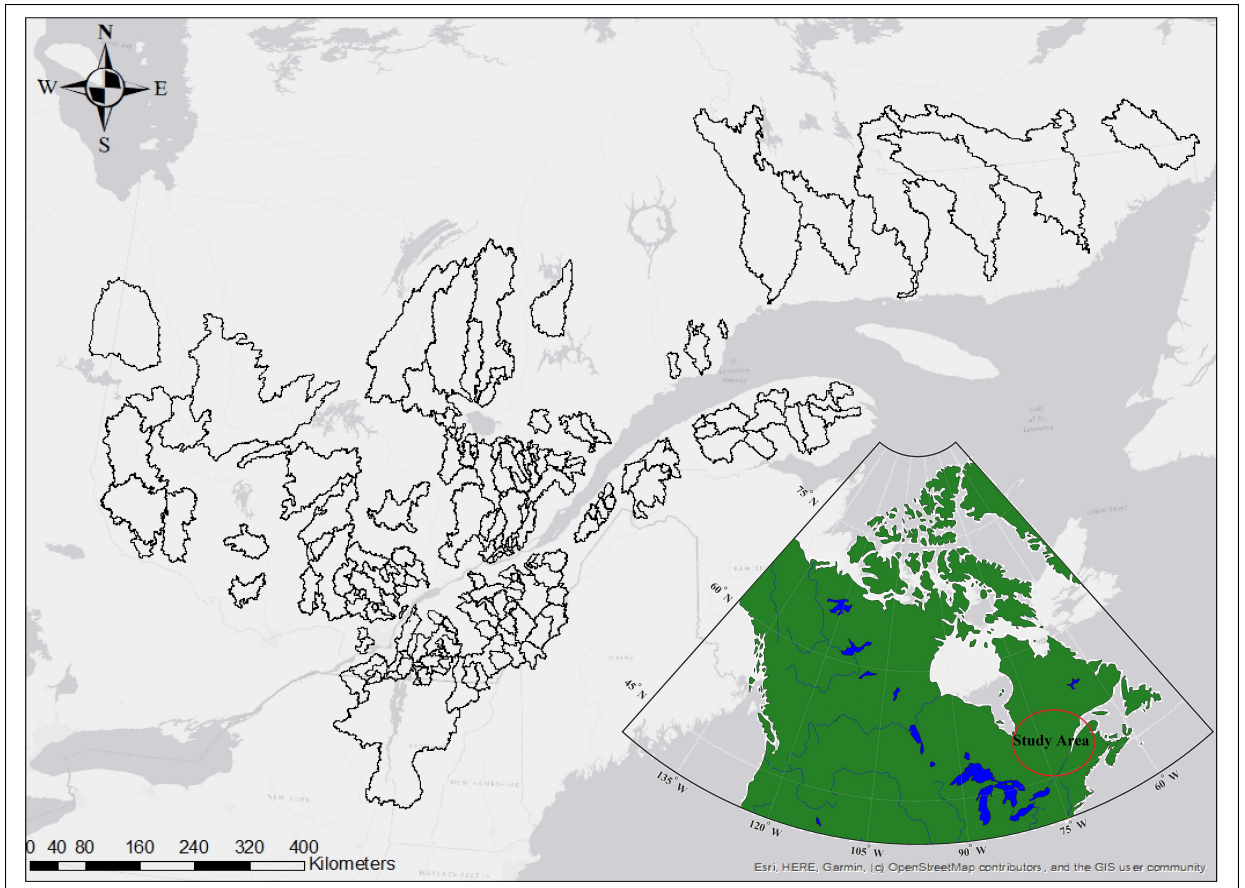
Figure 5.2   Distribution of the 171 catchments used in this study, located in the southern part of Quebec

for Hydrotel to assess the accuracy of parameter estimation at « pretend » ungauged locations. By « pretend » ungauged locations, we mean that even though those catchments are gauged, we pretend that they are not, so as to verify the capacity of the parameters estimated using the RF model to provide accurate streamflow simulations with Hydrotel. Finally, we use the regionalization model in the specific context of nested catchments to evaluate model efficiency for internal ungauged locations and examine hypotheses 2 and 3 according to the second column of Figure 5.3: Test and Analyses. We discuss the details of the procedures in the following.
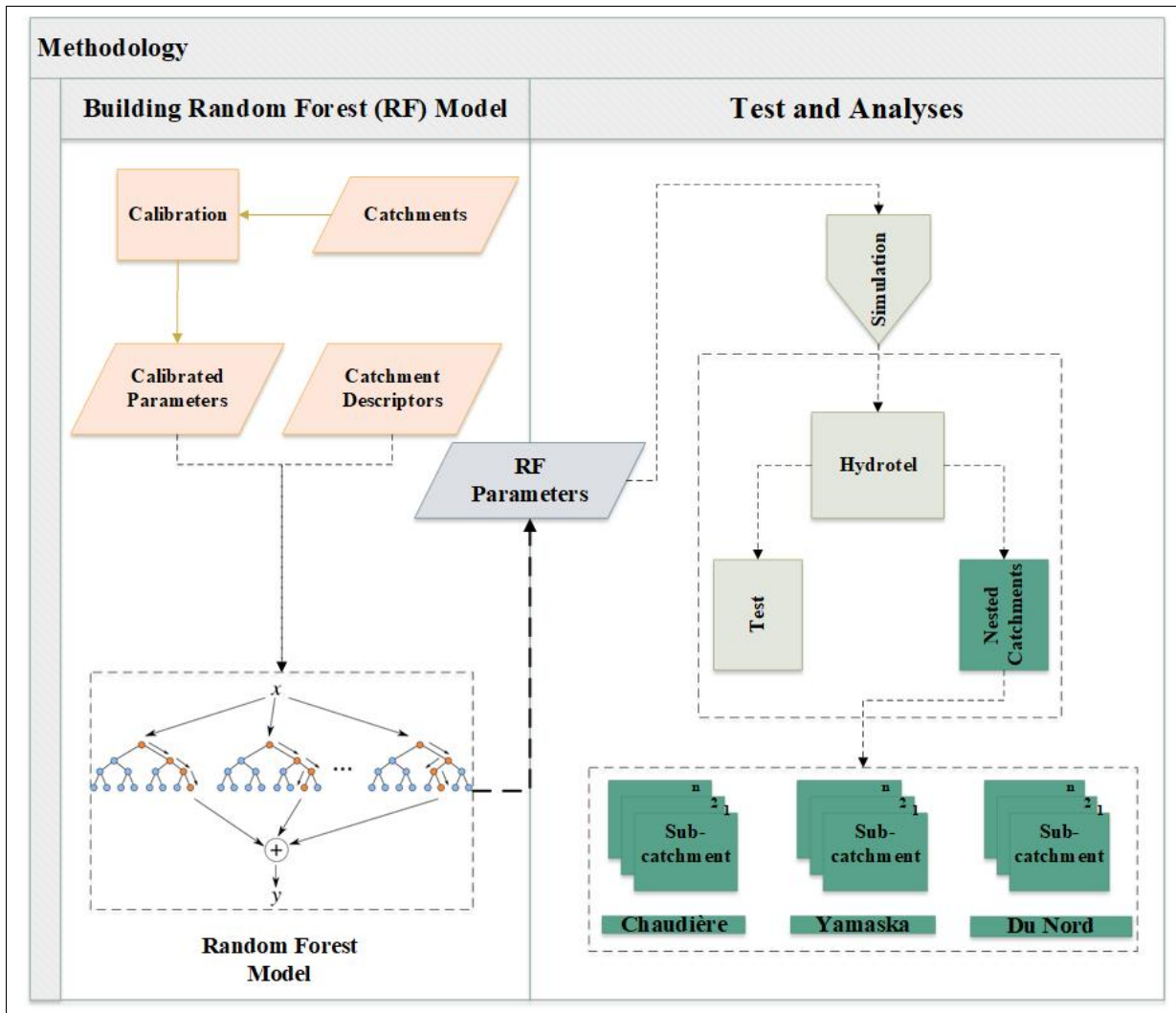
Figure 5.3   Schematic description of the methodology for this project

5.2.4   Calibration

The dynamically dimensioned search algorithm (DDS; Tolson & Shoemaker, 2007) is used to calibrate Hydrotel. Given that the procedure is computationally intensive, a set of 500 trials per catchment using the Kling-Gupta Efficiency (KGE; Gupta *et al.*, 2009) as the objective function is fixed. The KGE was computed using Equation 5.1

$$KGE = \sqrt{(r-1)^2 + (\frac{\sigma_{sim}}{\sigma_{obs}} - 1)^2 + (\frac{\mu_{sim}}{\mu_{obs}} - 1)^2} \qquad (5.1)$$

where r is the linear correlation between observations and simulations, σ_{sim} and μ_{sim} are respectively the standard deviation and mean of the simulated streamflows, whereas σ_{obs} and μ_{obs} are the standard deviation and mean of the observed streamflows.

The length of the calibration period varies between catchment depending on data availability. For most catchments, this period covers 10 years, between 1990 and 2017 in line with the available meteorological data (normally 2008-2017 unless the measurement of the streamflow stopped before 2017). Yet there are few catchments for which either measurements stopped before 2000 or have recently resumed and therefore have less than 10 years of data. The mean (median) KGE of calibration for 24-hour time-step is 0.81 (0.87) and for 3-hour time-step is 0.71 (0.84).

5.2.5   Catchment Descriptors

We define a set of catchment descriptors based on the existing literature (e.g., Arsenault *et al.*, 2014; Merz & Blöschl, 2004; Merz *et al.*, 2020). These catchment descriptors can be classified into four major groups. Table 5.2 presents the complete list of CDs used in this study along with their description, units, and range among the 171 catchments.

5.2.6   Random Forest Model

The Random Forest (RF) algorithm is derived from decision trees. A decision tree model divides the input space into a number of simple segments, typically using sequential binary decision rules. Each decision splits the region into two (or more) nodes which are referred to as leaves. The number of decision rules is referred to as the depth of a tree. The average (in case of regression problems) or majority of votes (in case of classification problems) for each segment determines the output of the model. Decision trees are intuitive, fast, and interpretable (Bishop & Nasrabadi, 2006). But there are issues such as greediness (a tree getting trapped in local optimum) and arbitrariness (sensitivity over early splits and over the data points) of the algorithms that limit their applications (James, Witten, Hastie & Tibshirani, 2013).

Table 5.2   List of catchment descriptors (CDs), their definition, group type and range (minimum, average, and maximum)

CD	definition	group type	min	average	maximum
AI	aridity index (ETPM/PRM)	meteorological	0.4	0.5	0.6
ETPM	mean annual potential evaporation[mm]	meteorological	403.8	528.7	640.5
PRM	mean annual precipitation [mm]	meteorological	886.6	1076.5	1372.5
TMM	mean annual temperature [$^\circ C$]	meteorological	0.3	3.6	6.8
TMAX	mean annual maximum temperature [$^\circ C$]	meteorological	1.6	4.9	8.1
TMIN	mean annual minimum temperature [$^\circ C$]	meteorological	-1.1	2.2	5.5
WR	water (% of surface area)	land cover	0.0	0.0	0.1
BS	bare soil (% of surface area)	land cover	0.0	0.0	0.2
DF	deciduous forest (% of surface area)	land cover	0.1	0.3	0.6
AL	agriculture lands (% of surface area)	land cover	0.0	0.2	0.9
CF	coniferous forest (% of surface area)	land cover	0.0	0.4	0.8
IR	impermeable (% of surface area)	land cover	0.0	0.0	0.8
BL	bogland (% of surface area)	land cover	0.0	0.0	0.4
WL	wetland (% of surface area)	land cover	0.0	0.0	0.2
SN	sand (% of surface area)	soil type	0.0	0.5	1.0
SL	sandy loam (% of surface area)'	soil type	0.0	0.4	1.0
SiL	silt loam (% of surface area)	soil type	0.0	0.0	0.7
CL	clay loam (% of surface area)	soil type	0.0	0.0	1.0
SA	surface area [km^2]	topographical	2.8	1800.0	22113.1
DD	drainage density [km/km^2]	topographical	0.0	0.3	2.3
LR	lake ratio[km/km^2]	topographical	0.0	0.0	0.2
WTI	wetness index	topographical	7.1	8.6	11.0
MSLP	mean slope	topographical	0.5	6.8	21.3
MEL	mean elevation	topographical	20.4	344.7	858.4
CVEL	lcoefficient of variation of elevation	topographical	90.0	186.8	305.0
MASP	median aspect	topographical	7.1	8.6	11.0
LAT	latitude [decimal degree]	topographical	44.9	47.2	52.2
LON	longitude [decimal degree]	topographical	-79.3	-71.6	-57.9

RF has been proposed to integrate out the effect of arbitrariness of decision trees. A RF is a collection of decision trees and can handle both regression and classification problems. The idea is to train multiple decision trees over multiple samples of the data and take the expectation of that ensemble as the output. In order to create multiple samples of data, bagging, which is a combination of bootstrap and aggregation, is used (Breiman, 2001). RF typically limits the maximum number of input variables corresponding to individual decision trees to maximize the variance of trees. Averaging over outputs of all trees therefore leads to reduction of the

error related to the variance of models. For this research a regression RF model is employed to approximate the hydrological model parameters

One drawback of ensemble tree-based models relative to a simple decision tree is the reduction of interpretability. While the splitting procedure in decision trees is transparent and can demonstrate the effect of each split in the space of input variables, it is difficult to interpret an ensemble of trees. This, however, can be rectified by computing the relative importance of input variables. We calculate the relative importance of input variables by recording the reduction of error corresponding with each input variable of a decision tree and average them out over all decision trees. As a result, a summary of the effect of input variables to approximate the model parameters can be obtained. Section 5.3.2 details the relative importance of input variables (here CDs) and the interpretability of the model.

5.2.7   Test and Analyses

For training the model (we use training for RF model and calibration for Hydrotel to distinguish the two), the catchments with poor KGE in calibration are filtered out to guarantee the quality of the training data. Furthermore, the catchments are randomly split into training and testing datasets to evaluate the efficiency of the RF model. In the next step, the efficiency of regionalization for both 24-hour and 3-hour time-steps are obtained for the test dataset to evaluate the first hypothesis. The relative gap between the Hydrotel calibrated parameters and RF approximated parameters is calculated by Equation 5.2. This measure helps compare the efficiency of the simulations with 3- and 24-hour time-steps for validation of the first hypothesis. The comparisons are not limited to this measure (Equation.5.2) and other statistical indices (mean, median, variance of the distributions) are used for evaluation of the model efficiencies. Moreover, the two samples t-test is employed to evaluate the statistical significance of the differences between 3- and 24-hour RF and calibrated simulations.

$$Relative\ Gap = \frac{(median\ of\ calibration\ efficiencies) - (median\ of\ RF\ efficiencies)}{median\ of\ calibration\ efficiency}$$

$$(5.2)$$

Finally, the 2nd and 3rd hypotheses are examined through the framework of nested catchments. A nested catchment is a parent catchment, which is normally large, that comprises multiple smaller catchments (or sub-catchments hereafter). The streamflow at different points inside the catchment is measured. Therefore, it is possible to evaluate the regionalization technique at the internal ungauged locations. There are three such catchments (i.e. Chaudière, Yamaska, and du Nord) in the study areas.

Figure 5.4 and 5.5show the DEM and land-cover map of the (a) Yamaska, (b) du Nord, and (c) Chaudière catchments. The Yamaska catchment includes five sub-catchments as can be seen in Figure 5.4. In terms of topographic features, the catchment on average is mostly flat, except its relatively moderate slope on the eastern side. In terms of land-cover (Figure 5.5 a), the catchment is diverse: The western part of the catchment is covered by agricultural lands while deciduous forest is the main feature of the central and eastern parts. There is also a rather large lake in the forested area.

The du Nord catchment includes three sub-catchments. The maximum elevation is even lower than for the Yamaska catchment, but the proportion of high altitude pixels (Figure 5.4) is higher. Also, while the dominant land cover feature (Figure 5.5 b) is deciduous forest, a network of coniferous, lake, impermeable, and agricultural areas give a degree of diversity to this catchment.

Chaudièreis is the parent catchment of two sub-catchments. The DEM of the catchment (Figure 5.4) shows the topography that is mostly flat except the steeped southern and eastern edges. The dominant feature of land-cover (Figure 5.5 c) is forest, which is a mix of coniferous and deciduous trees. Agricultural lands are also scattered across the central part and outlet of the catchment. Megantic Lake can also be seen in the southern part.

We approximate the parameters with the RF model at different levels of spatial discretization for nested catchments (i.e catchment level, sub-catchment level, and RHHU level) and investigate their spatial consistency when the spatial resolution is refined (See section 5.3.3). To investigate hypothesis II, the catchment descriptors at the RHHU level of discretization are extracted from the nested catchments and plugged into the RF model to approximate parameters per RHHU.
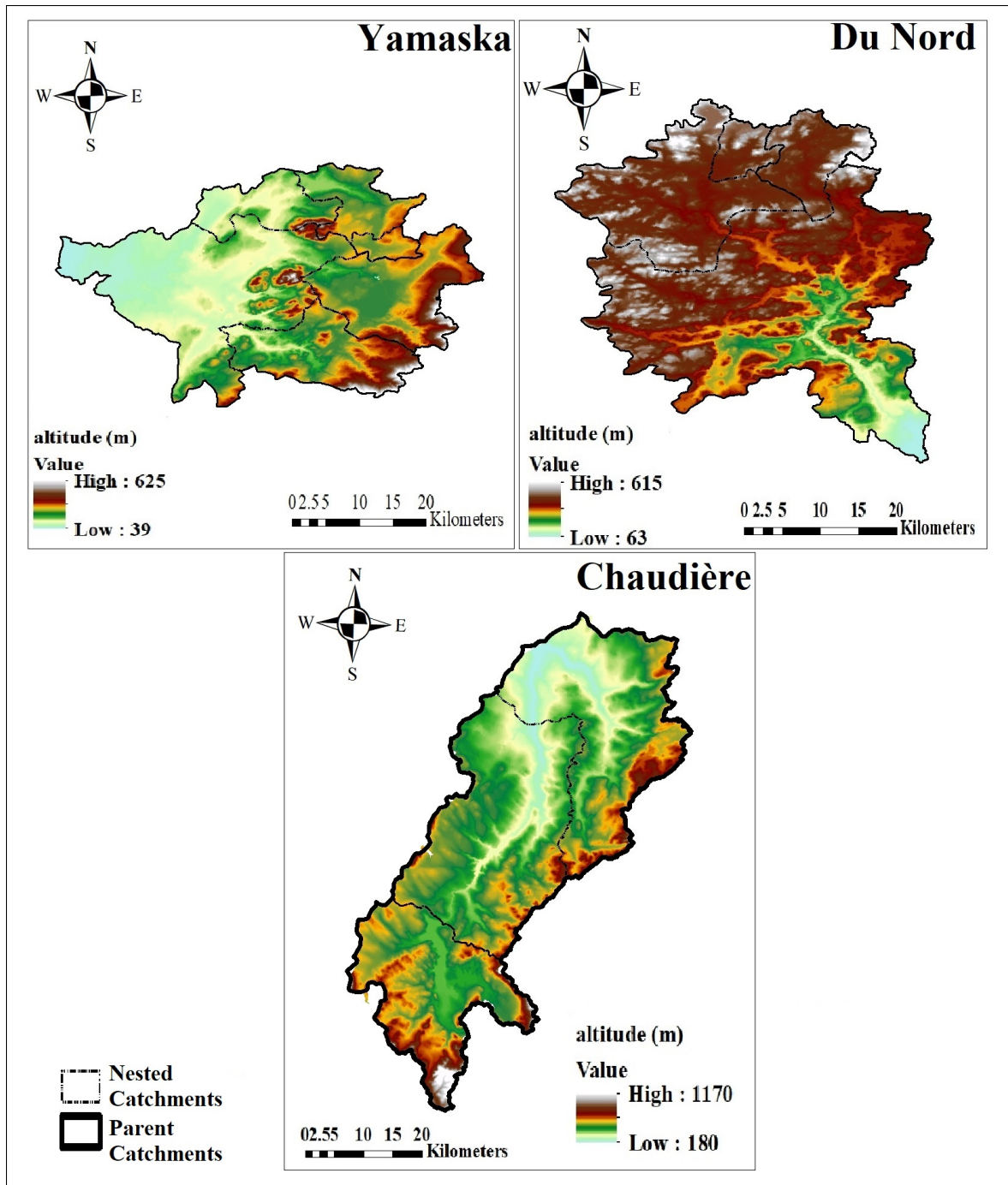
Figure 5.4 Digital elevation model and boundary of the nested catchments

Subsequently, the spatial consistency of the approximated parameters at the RHHU level is evaluated by calculating the spatial correlation of the mean elevation (an example CD) and two
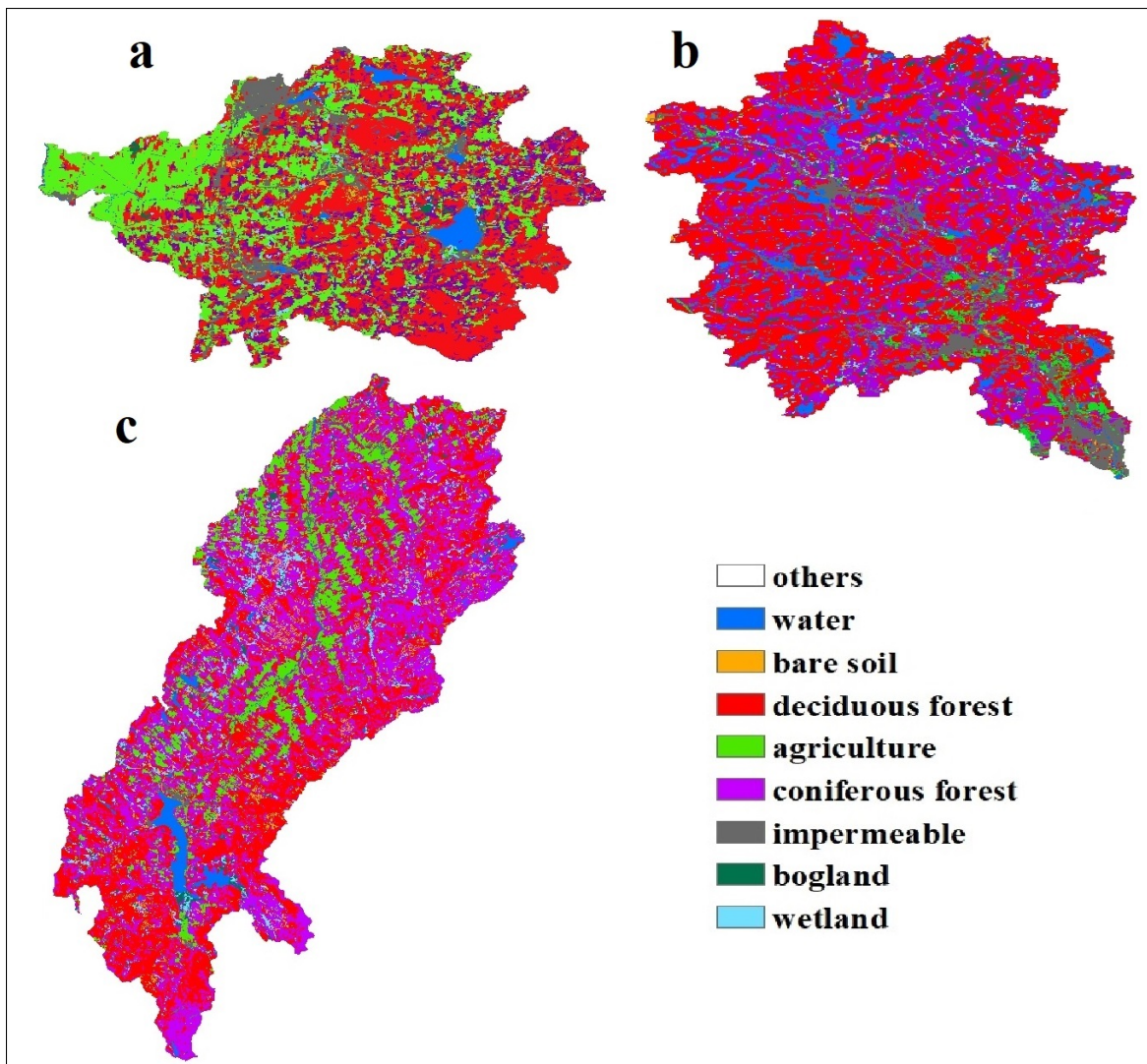
Figure 5.5   land use map of the nested catchments. a: Yamaska, b: Du Nord, c: Chaudière

model parameters: L1 and MTN. Spearman correlation coefficient and least squares regression are computed. These two criteria are further calculated for the catchment discretization level and compared with the RHHU level to investigate if transferring parameters from the catchment to RHHU discretization levels reproduces similar patterns. Note that, for calculating the spatial correlation at the catchment level, the catchments included in the test dataset are used. Moreover, a visual analysis of the parameters' distribution at the RHHU and sub-catchment levels is performed to highlight the hydrological realism of the parameter approximation at different

levels of discretization (For this point we performed the tests on Yamaska, as this catchment shows diversity according to Figure 5.5).

To investigate hypothesis III, we performed simulations with Hydrotel using the parameter sets corresponding with different levels of discretization derived by RF for the nested catchments (See section 5.3.4). Inline with the methodology devised to verify hypothesis II, we define three levels of simulations. 1- Simulations with fully-distributed parameters (FDP), which corresponds with approximation of parameters by RF at the RHHU level; 2- simulations with semi-distributed parameters (SDP), which corresponds with approximation of parameters by RF at the sub-catchment level; and 3- simulations with lumped parameters (LP), which corresponds with approximation of the parameters by RF at the catchment level. The average efficiency of these simulations over subcatchments are further compared for 3- and 24-hour time-steps to highlight the sensitivity of the model over varying spatio-temporal discretization.

5.3   Results and Discussion

In this section the results are presented following this order: Section 5.3.1 provides the RF model evaluation for 3- and 24-hour time-steps and compares it with the calibration results to investigate hypothesis I. Section 5.3.2 evaluates the relative importance of CDs in the determination of model parameters. Section 5.3.3 explores parameter transferability across scales to verify/refute hypothesis II. Section 5.3.4 is dedicated to multiscale hydrological simulations and sensitivity of the hydrological model to the spatial discretization to validate hypothesis III.

5.3.1   Model Evaluation

Figure 5.6 illustrates the results of simulations carried out with regionalized (RF) and calibrated (Cal) parameters for 3- and 24-hour time-steps. Subplot a compares the empirical cumulative distribution function (ECDF) of the modeling efficiencies. A summary of the statistics regarding the ECDFs can be found in Table 5.3. The mean (median) of RF distributions for 24- and 3- hour time-steps are 0.68 (0.7) and 0.75 (0.76) respectively, evidencing an improvement of

the simulations when a finer time-step is used. The 24-hour time-step training data has an advantage over the 3-hour time-step for calibration in terms of efficiency: the mean (median) of distributions for 24- and 3-hour time-steps are 0.88 (0.9) and 0.84 (0.85) respectively. As expected from any regionalization techniques, there is a loss of efficiency from calibration to regionalization. However, this relative gap (Eq 2) is considerably larger for 24-hour time-step simulations (22%) than for the 3-hour time-step (12%). Furthermore, the standard deviation (std) of 24-hour RF distributions is larger (std=0.12) than for 3-hour (std=0.07) time-step.

Figure 5.6 subplot b shows the distributions of RF and calibration parameters in 3- and 24-hour time-steps. Here, the RF and calibrated parameters are jointly standardized between 0 and 1 to facilitate the comparisons. As can be seen, the median of the RF and calibration distributions are in close proximity for most of the cases particularly for mixed degree-day-energy balance parameters (e.g. MTD to MRC), which demonstrates models successfully approximated the median of the parameter distribution. However, the spread of calibration parameters is larger than that of RF parameters and it seems that the regionalization technique tends to systematically underestimate the spread of the parameters. Comparing the time-step of simulations, the 3-hour RF parameters (i.e. RF-3h) have a better approximation in terms of the median of the corresponding calibration parameters (i.e. Cal-3h) than that of 24-hour time-step parameters (e.g. the distance between the median of RF and calibration parameters for L1, L2, MRD, MRN, and TSR is larger for 24-hour than that for 3-hour time-step) results in better simulations, which is inline with the first hypothesis.

To explore the statistical significance of the differences between 24-hour and 3-hour simulations, a two-sample t-test was used. In this test, the null hypothesis is that input vectors are independently sampled from the same distribution (if input vectors are from different distributions the test returns 1). The test was applied to the ECDFs of calibration (i.e. Cal-3h and Cal-24h in Figure 5.6 subplot a) and also to the ECDFs of RF (i.e. RF-3h and RF-24h in Figure 5.6 subplot a). The goal is to investigate the existance of statically significant differences between ECDFs, which could be attributed to RF simulations when the time-step changes. The results show that while there are not any statistically significant differences regarding calibration-ECDFs, a significant
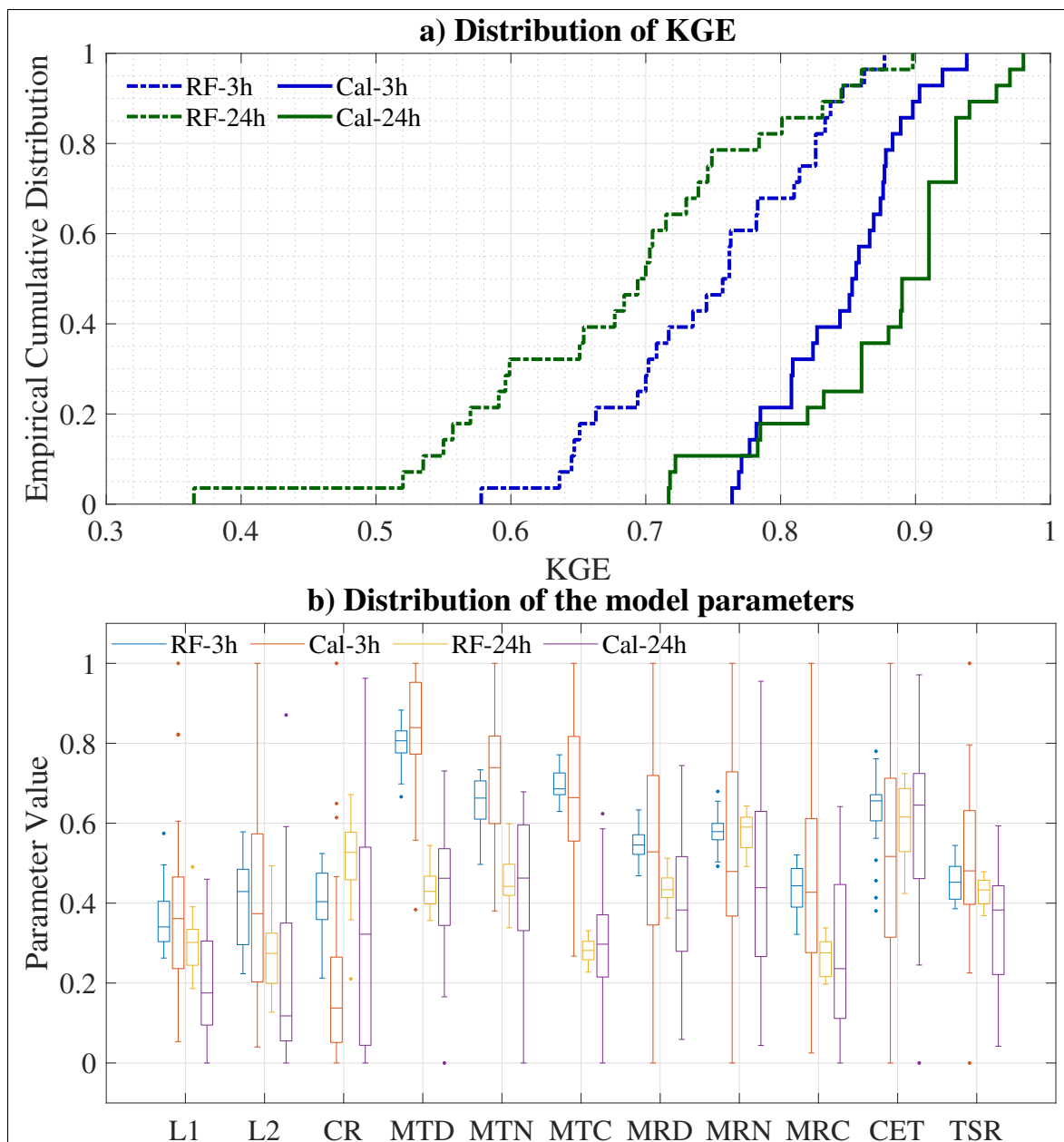
Figure 5.6 Comparing calibration and regionalization simulations for the test dataset (RF=Random Forest, Cal=Calibration). a) Distribution of the regionalization and calibration KGE for 3-hour and 24-hour time-steps. b) Standardized distribution of calibrated and approximated parameters for 3- and 24-hour time-steps respectively

difference at 5% level can be found in RF-ECDFs when the time-step of the simulations varies from daily to sub-daily (Hypothesis I is verified for our case study).

Table 5.3   The modeling efficiency (KGE values)
statistics calculated for the test dataset in 3- and 24-hour
time-steps (RF=Random Forest, Cal=Calibration).

Statistics	RF-3h	RF-24h	Cal-3h	Cal-24h
Mean	0.75	0.68	0.84	0.88
Median	0.76	0.70	0.85	0.90
Standard deviation	0.07	0.12	0.05	0.07
Null hypothesis (t-test)	1		0	

5.3.2   Relative Importance of CDs

Figure 5.7 demonstrates the relative importance of CDs to approximate two selected parameters (L1 and MTN) for 24-hour and 3-hour time-steps respectively. Those parameters are considered to be representatives of the vertical water budget and snow melt/accumulation in Hydrotel. As can be seen, no distinct group of features (i.e. meteorological CDs, topographic CDs, soil type CDs, land use CDs) controls parameter approximation. For both parameters, the groups of meteorological (first six, AI to TMIN) and topographic (last 10, SA to LON) data have similar weights. This is expected as these two groups should be correlated with one another (Merz & Blöschl, 2004; Merz *et al.*, 2020). Furthermore, Figure 5.7 shows that land-cover features (WR to WL) are important, particularly for approximating MTN, which is responsible for snow accumulation and melt. The soil type (SN to CL) shows a lesser degree of importance to determine the parameters. This might be due to the rather uniform soil texture of the study area (based on Table 5.2, the median of SiL and CL is zero, showing that the soil type of the catchments is rather homogeneous with the dominance of sand and sandy loam). The most important individual features are elevation-related input variables (MEL, CVEL, MASP), mean annual precipitation and temperature (PRM, TMM), and the percentage of coniferous or deciduous forests (CF, DF).

According to Figure 5.7, no systematic pattern relating CDs to parameters can be found. In fact, the weights of many CDs are unexpectedly high in relation to some parameter (e.g. LON to determine L2). This might be related to equifinality and identifiability in the calibration process.
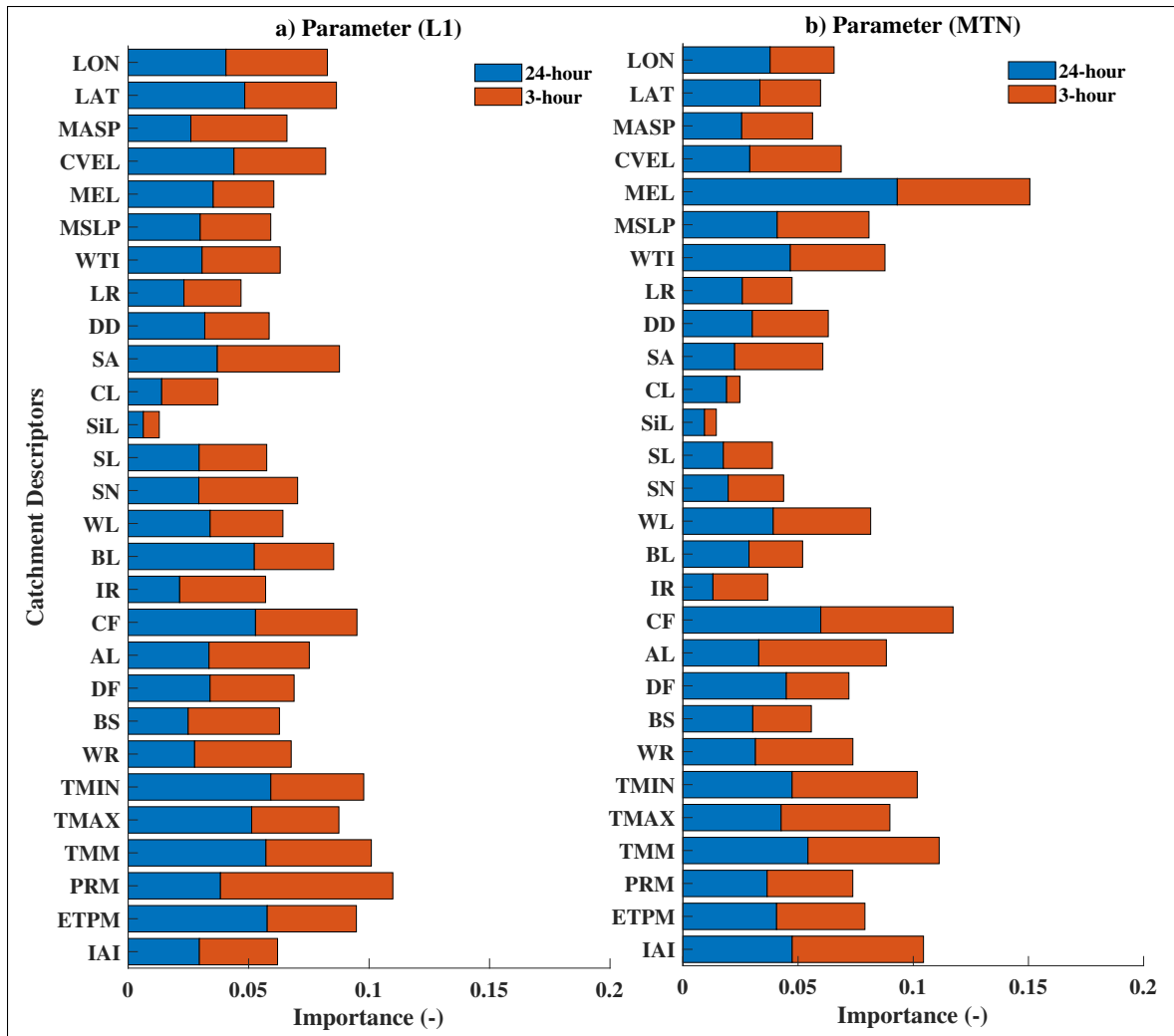
Figure 5.7   Relative importance of predictor features in parameters approximated by
RF for simulations with 3-hour time-step

Indeed, it is possible that there exists more than one good parameter set for each catchment, in particular when model calibration is based on a single type of observation (streamflow) which might not be enough to constrain the optimization problem. If so, the parameter set used for a given catchment in the training of the RF model is nothing but a possibility from a sample of many other possibilities, not to mention the possible interrelations between the model parameters. This of course could induce uncertainty in the training of the RF model. Furthermore, the multivariate nature of calibration may cause an inter-correlation of parameters. Therefore, it is expected that a CD that controls a specific parameter affects many unrelated parameters due to

124

the correlation between parameters. These observations support the conclusion from Section 5.1 that it is difficult or improbable to constrain the parameters to an a priori relationship to be used as a transfer function because some information may remain hidden to the system and reduce the performance of the regionalization model.

5.3.3   Transferring Parameters Across Scales

To test the transferability of RF approximated parameters from the catchment scale to the RHHU, we explore the existence of consistent spatial correlation between CDs and parameters across different levels of discretization. Accordingly, the spatial correlation between the selected parameters in Section 5.3.2 (i.e. L1 and MTN) and mean elevation was calculated. Mean elevation was selected, as it is among the leading predictor features controlling approximation of RF parameters according to Figure 5.7.

Figure 5.8shows the results for the spatial correlation between mean elevation and RF approximated parameters at different levels of spatial discretization. Figure 5.8 a to e presents the spatial correlation between parameters and mean elevation for the test dataset with 24- and 3-hour time-steps at the catchment level (CL) (for CL resolution, catchments belonging to the test dataset as discussed in Section 5.8 were used). Figure 5.8 e to p shows the spatial correlation between parameters and mean elevation at the RHHU level (RL) for three selected catchments (Chaudière, Yamaska, and Du Nord) as discussed in Section 5.2.

Figure 5.8 a shows that the L1 parameter is positively correlated to the mean elevation (cr=0.5) at the catchment level for the 24-hour time-step. Similarly, Figure 5.8 e, i, and m, shows positive correlation coefficient at the RHHU level, and the values are similar to those obtained at the catchment level. Particularely, the Yamaska and Du Nord catchments (Figure 5.8 i and m) show higher degree of similarity to the catchment level (Figure 5.8 a) in terms of the magnitude of correlation coefficient and the least squares regression line (drawn in red). Regarding the 3-hour time-step for L1 (Figure 5.8 b, f, j, n), there are larger differences in the magnitude and even direction of the correlation coefficient between catchment and RHHU levels than that of 24-hour
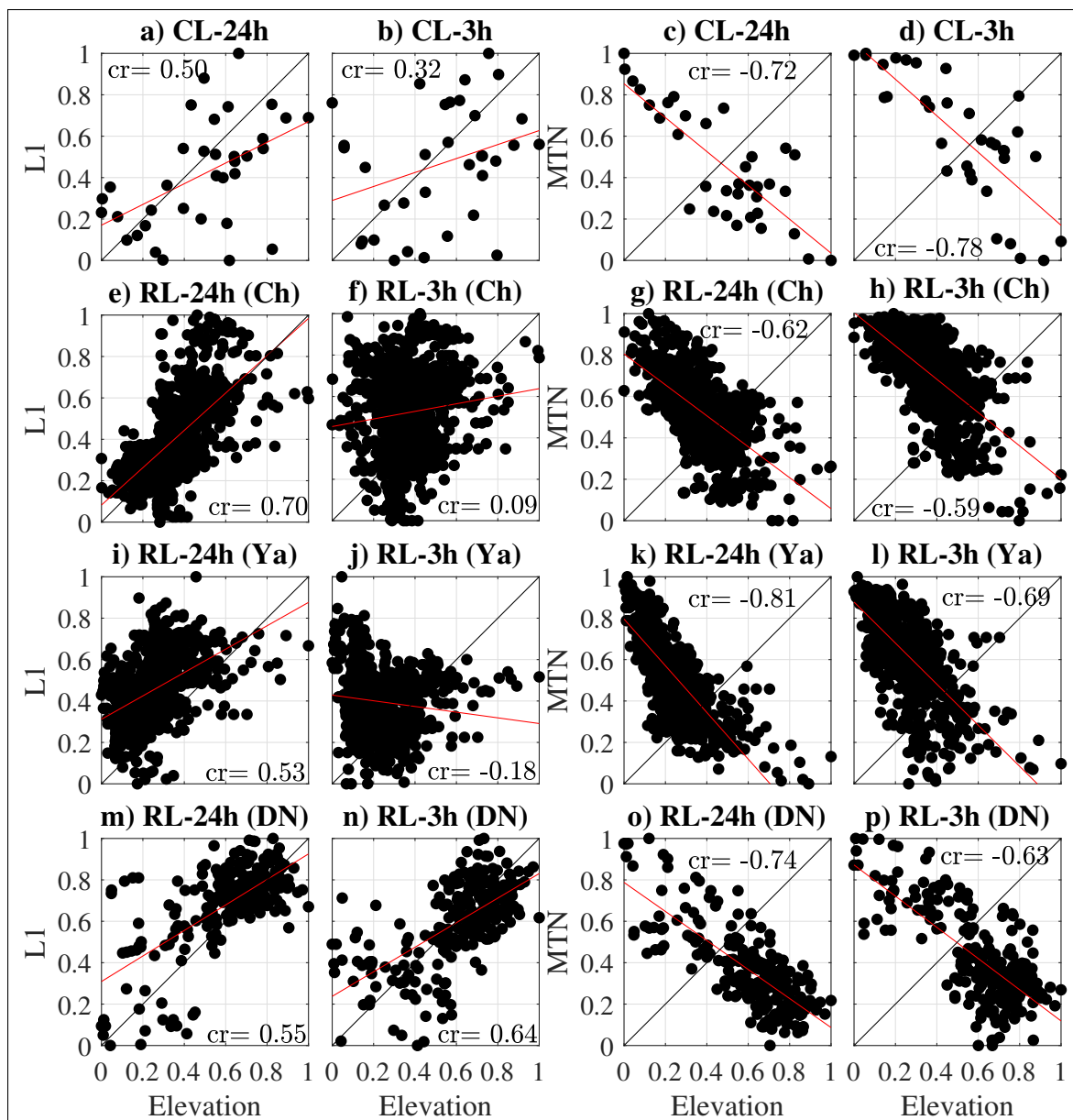
Figure 5.8   Spatial correlation of RF approximated parameters (L1 :first layer thickness and MTN : melt threshold non-forest) and elevation: a to d- Catchment Level (CL) parameter resolution; e to p- RHHU Level (RL) parameter resolution. Ch, Ya, and DN represent Chaudière, Yamaska, and Du Nord catchments respectively. The red line shows the least squares regression line

time-step. Interestingly, the spatial correlation at the catchment level is 0.32 in Figure 5.8 b, while Yamaka shows a 0.18 correlation coefficient in the opposite direction according to Figure

5.8 j. This might be due to the rather lower importance of mean elevation as CD for determining L1 at the 3-hour time-step than that of the 24-hour time-step according to Figure 5.7.

Figure 5.8 c, d, g, h, k, l, o, p shows that there is a high negative correlation between MTN and mean elevation regardless of the level of discretization, catchment, or time-step of the simulations. Also, the least squares regression line is similar for all cases. The reason for such similarity is the high degree of importance of the mean elevation (MEL) to determine parameters, as evidenced in Figure 5.7.

To further explore the transferability of parameters across scales, Figures 5.9 illustrates the distribution of RF-estimated parameters at the 24-hour time-step (L1 and MTN) across the Yamaska catchment at the RHHU and sub-catchment levels. Comparing Figure 5.9 a and c with Figure 5.5 a, there is an impact of the presence of agricultural lands on the distribution of the parameters. Notably, the distinct increase in the magnitude of L1 at the western side of the catchment (Figure 5.9a) matches the agricultural area in Figure 5.5 a, evincing a direct link between land cover and parameter distribution at the RHHU discretization level. Yet the effect of agricultural land on the L1 distribution at the sub-catchment level (Figure 5.9 b) cannot be directly observed due to lower resolution of the simulation than that for RHHU level, which leads to the spatially averaged values for L1 over the sub-catchments.

A similar pattern is visible regarding the MTN distribution: a smooth transition from west to east at the RHHU level (Figure 5.9 c), in contrast with an abrupt transition of MTN in the eastern side of the catchment at the sub-catchment level (Figure 5.9 d). In panels c and d, the distribution of MTN appears mainly controlled by the distribution of elevation (i.e. higher elevation leads to a lower MTN) according to Figure 5.4 (Yamaska), which is consistent with Figures 5.7 (i.e. high degree of importance for mean elevation) and 8 k (i.e. high negative correlation between MTN and mean elevation).

Figures 5.8 and 5.9 show that the distribution of parameters at the RHHU level is spatially consistent, in the sense that a spatial correlation between variation of CDs and parameters is observable. Therefore, RF-approximated parameters can reasonably reflect the characteristics of
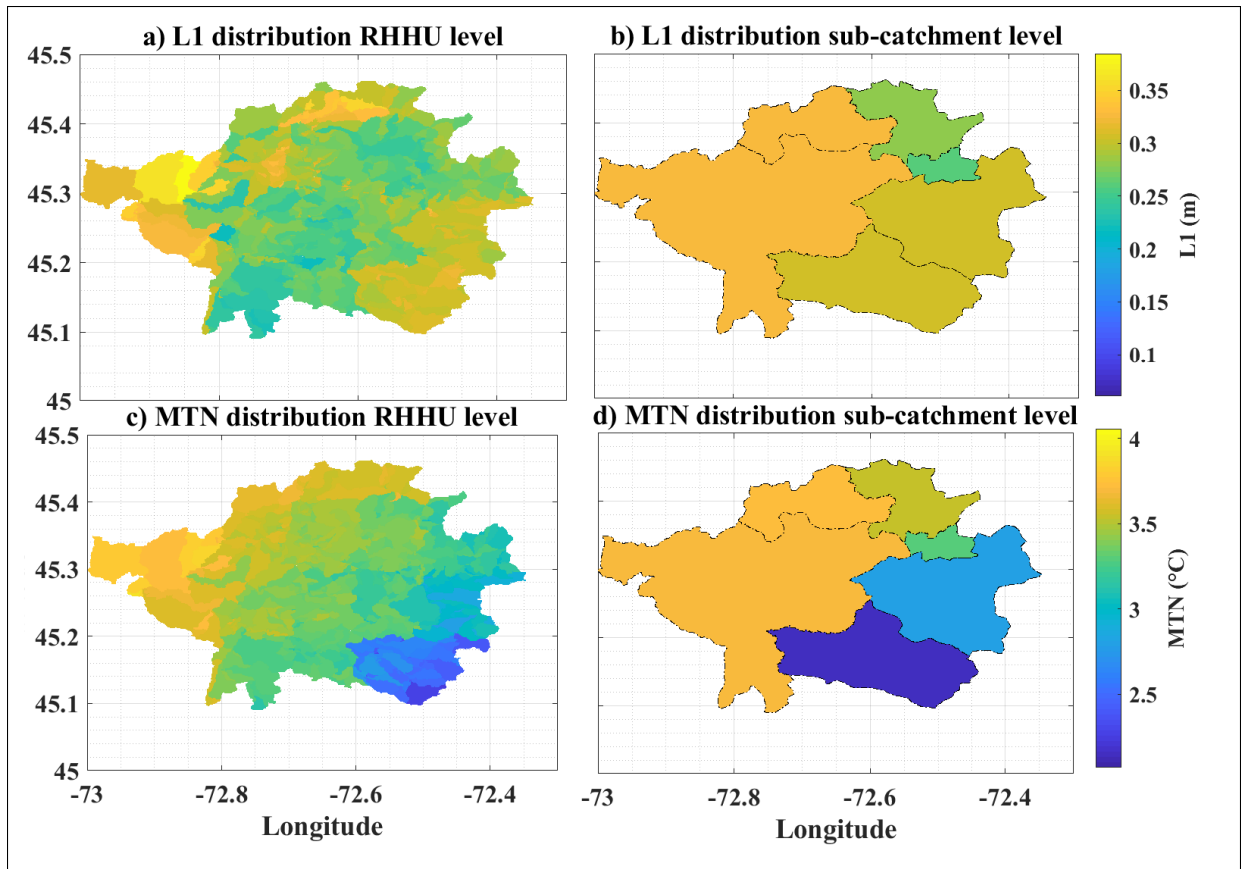
Figure 5.9   Distribution of RF approximated parameters (L1 and MTN) in 24-hour time-step across the Yamaska catchment at RHHU (subplots a and c) and sub-catchment (subplots b and d) discretization levels

the catchments at finer resolutions even if the resolution of the training data is coarse (Hypothesis II). In the next section, we will verify if using more spatially representative CDs has any benefit in terms of efficiency of the regionalization scheme (Hypothesis III).

5.3.4   Multi-Scale Simulation

Figure 5.10 shows the average KGE of all nested catchments for different discretization levels. Figure 5.10 a and b show the mean KGE of simulations for Chaudière subcatchment; Figure 5.10 c and d show the mean KGE of simulations for the Yamaska subcatchments; and Figure 5.10 Subplots e and f show the mean KGE of simulations for the du Nord subcatchments, all for 3-hour

and 24-hour time-steps respectively. The average efficiencies show consistent improvement as the number of distributed parameters increases at the 24-hour time-step (Figure 5.10 a, c, and e). The efficiency for Yamaska (Figure 5.10 c) shows the greatest improvement (12%) followed by du Nord (6%- Figure 5.10 e) and Chaudière (4%- Figure 5.10 a). This is expected as Yamaska is the most heterogeneous catchment, with a distinguishable transition from forest to agricultural lands (Hypothesis III is verified for those catchments at the 24-hour time step). Therefore, simulations with fully distributed parameters helps better represent catchment heterogeneity leading to an increase in model skill for the Yamaska catchment .
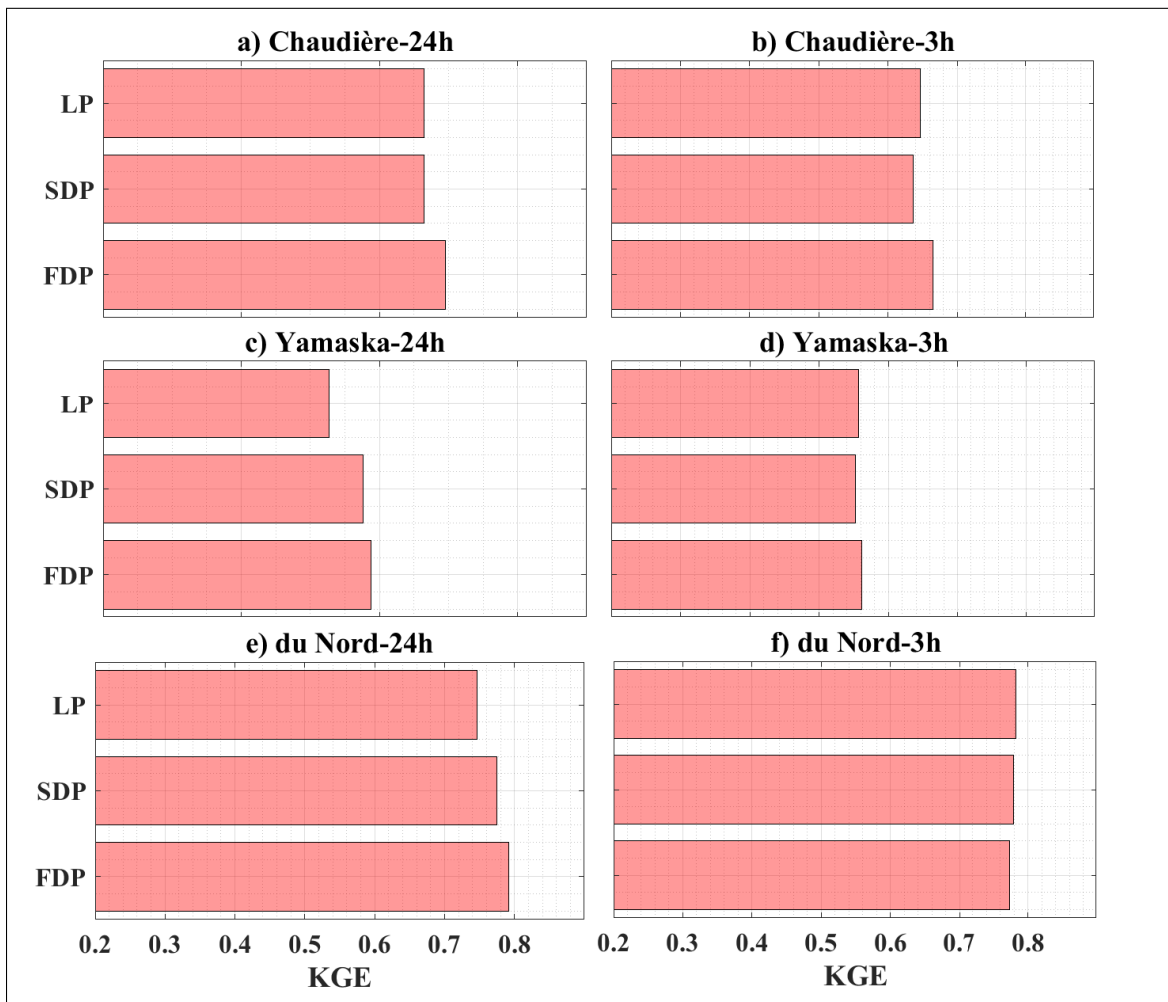
Figure 5.10   Average efficiency of RF simulations for each nested catchment at different levels of parameter discretization: Fully Distributed Parameters (FDP), Semi Distributed Parameters (SDP), and Lumped Parameters (LP)

No pattern can be detected regarding 3-hour time-step simulations, even for the most diverse catchment in terms of land cover types (Yamaska). This observation is in line with the results presented in Figure 5.8, at least for the L1 parameter. The spatial correlations between this parameter and the mean elevation at different levels of the spatial discretization is less consistent for the 3-hour time-step than for the 24-hour time-step. The results suggest that such inconsistency in transferring L1 from the catchment to RHHU level is a potential cause for the lack of improvement in modeling efficiency at the 3-hour time-step when the spatial resolution is refined. This premise however cannot be verified here as there are multiple parameters and catchment descriptors involved in the estimation of parameters at different levels of discretization. In the following we analyze the hydrograph of the Yamaska catchment at 3- and 24-hour time-steps to investigate the impact of the choice of time step.

Figure 5.11 shows the mean annual streamflow at the outlet of the Yamaska catchment with 24- and 3-hour time-steps. Here, the streamflow is reproduced at different levels of discretization and compared with observed streamflow. For clarity, a segment of the hydrograph from mid-May to early June is selected and enlarged. Figure 5.11 a, shows a lag between simulations with lumped parameters (red) and simulations with semi- and fully-distributed parameters (green and blue respectively).Using spatially distributed parameters improves the simulation by removing this lag. For 3-hour simulations, the lag between the lumped parameters simulations, and the others has been considerably reduced due to the finer temporal resolution. Therefore, using RF-approximate parameters at the RHHU and sub-catchment levels has limited potential to improve the simulations when the temporal resolution is fine (Hypothesis III is not verified for this case study, with 3-hour simulations).
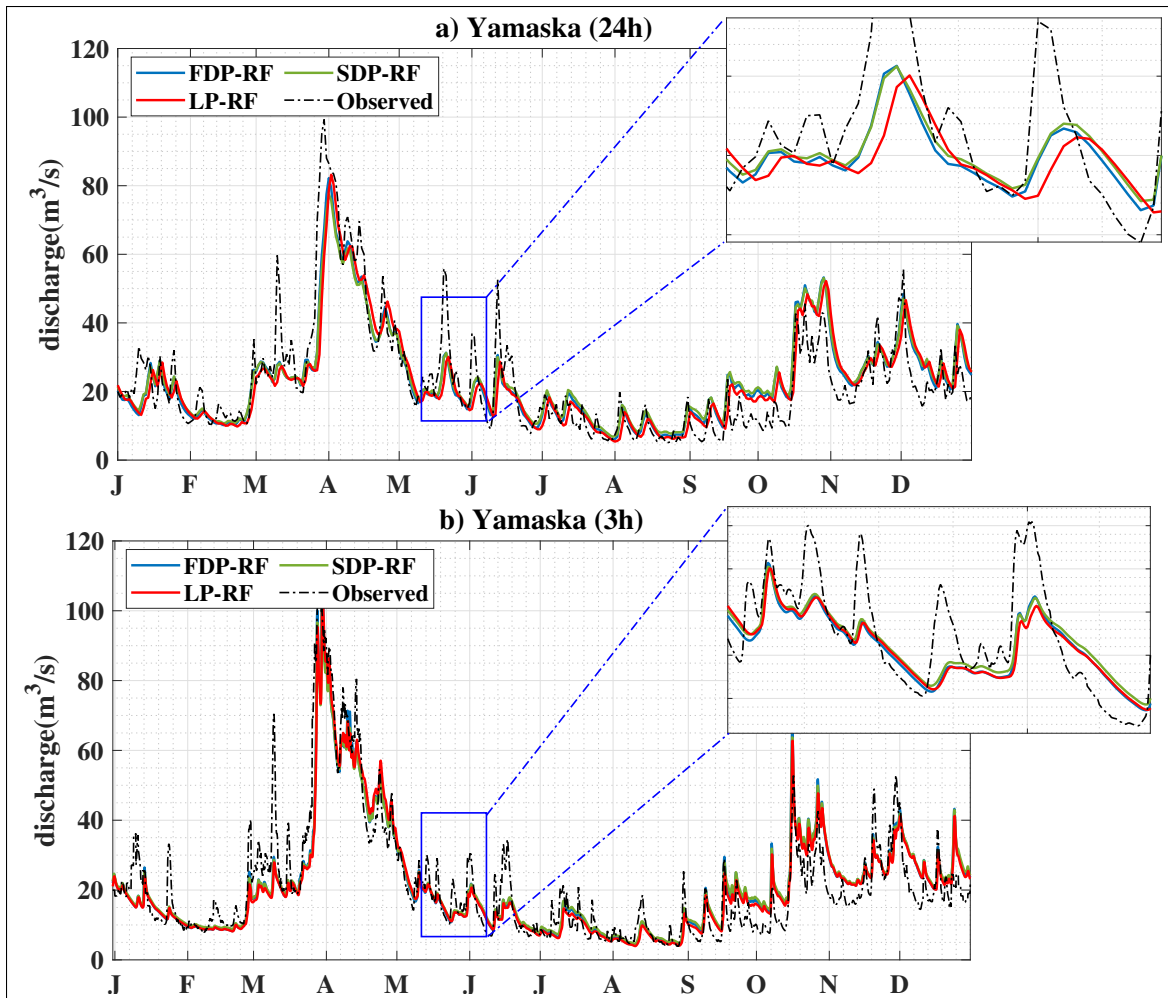
Figure 5.11   Mean annual hydrograph for the Yamaska catchment

5.4   Conclusion

In this paper, we used a ML technique (i.e., RF) in conjunction with a process-based hydrological model for regionalization. The approach showed skill in simulating streamflow in pseudo-ungauged catchments (Figure 5.6). We further ran the model across different spatio-temporal resolutions and investigated three hypotheses: (I) Finer time-step adds more information to the calibrated parameters and improves the efficiency of the regionalization model. (II) The parameters approximated by RF are spatially consistent and transferrable across the spatial

scales. (III) Refining the spatial resolution of CDs will improve the regionalization skill. The summary and conclusion about each hypothesis are the following:

1. Refining the time-step of the regionalization shows that simulations at a 3-hour time-step outperform a 24-hour time step (Figure 5.6). The two samples' t-test confirmed its statistical significance at a 5% significance level (table 5.3). Therefore, the results verify hypothesis I for this case study.

2. Approximating the parameters at different discretization levels (catchment, sub-catchment, and RHHU) shows that the parameters are spatially consistent. Figure 5.8 shows that the regionalization model maintains the spatial correlation between CDs and parameters. Furthermore, Figure 5.9 shows that the spatial distributions of the model parameters reflect the variation of topographic and land cover characteristics. Therefore, the results verify hypothesis II for this case study.

3. Figure 5.10 shows that using spatially refined CDs at daily time-step improves the regionalization skill, depending on the degree to which the catchment is heterogeneous (i.e., higher skill for more heterogeneous catchments). However, a clear pattern once the time-step is 3-hour has not been found (Figures 5.10 and 5.11). Hence, the results cannot verify hypothesis III.

CHAPTER 6

GENERAL DISCUSSION

This research suggests three general hypotheses defined under an umbrella objective, which is to quantify the impact of the spatio-temporal representation in distributed hydrological models. The hypotheses are as follows:

1. The choice of level of spatio-temporal discretization alters model parameters, which leads to uncertainty in streamflow and flood simulation. By increasing catchment area, the contribution of the choice of spatial scale and hydrological model in such uncertainty increases, and that of time-step decreases.

2. For small catchments ($< 500 km^2$), refining temporal resolution of simulation (from daily to subdaily) increases the relative change (from reference to future) of extreme summer-fall flow. Refining temporal resolution will not significantly affect projected extreme summer-fall flow for large catchments ($> 1000 km^2$).

3. There does exist relationships between model parameters and catchment descriptors that can be approximated by the random forest method. The underlying information related to catchment characteristics is transferable across scales through that approximation. Using fine-scale (in time and space) catchment descriptors improves the skill of the regionalization model.

This chapter provides a general discussion of the key findings of this research to link the results presented in chapters 3 to 5. Each hypothesis will be discussed individually. This is followed by specifying the study's limits, future work recommendations, and a brief summary.

6.1 Uncertainty linked to heterogeneity

Spatial resolution- The results presented in Chapters 3 and 4 suggest that there is uncertainty linked to the spatial resolution. This is in line with the studies focusing on peak flow simulation of urban catchments (Cao *et al.*, 2020a; Ichiba *et al.*, 2018; Warsta *et al.*, 2017). This uncertainty can be seen in simulation of low, high, and extreme flow (Figures 3.5, 3.11 to 3.13), extreme

flow projection (Figures 4.6 and 4.10), and other hydrological variables such as AET, snow depth, and base flow (Figures 3.6 and 3.7). This uncertainty has no correlation with catchment size (Figures 3.3,3.4,3.5). Several factors are included in the modeling behavior once the spatial resolution is altered. In two examples (catchments Aux Brochets and Boyer), using different spatial resolutions changed the spatial distribution of different hydrological variables and affects the final streamflow response at the outlet (See Appendix I and Section3.4). Such response might be a result of their relatively flat surfaces with a large portion of farmlands (see Table I-2).

Hydrological model- The choice of hydrological model is important in determining the spatio-temporal variability. Figure 3.5 shows a wider spread for WaSiM than that for Hydrotel. Moreover, Figure 4.10 shows that the choice of model contributes largely to the variability of simulated summer-fall extreme flow. For Hydrotel, grids are aggregated to create hydrological response units (HRU). Such aggregation can eliminate spatial information like surface connectivity and subsurface variability. Unlike Hydrotel, in WaSiM, the hydrological processes are computed at the scale of the raw spatial resolution of the inputs (i.e., land use matrix, DEM, and soil type matrix), without any aggregation to generate simulation units. As a result, spatial resolution variation creates a larger variability for simulations with WaSiM than for Hydrotel. While the response of more sophisticated process-based models is expected to be more realistic, there is no guarantee for a better simulation by these models (Seiller *et al.*, 2017; McDonnell *et al.*, 2007; Beven, 2000; Savenije, 2009; Clark *et al.*, 2011). Comparing WaSiM and Hydrotel, at least in streamflow simulations, the higher level of model complexity for WaSiM leads to a larger uncertainty corresponding to spatio-temporal discretization, with no significant benefit in terms of modeling skill.

Time-scale- Based on the results presented in chapter 3, the temporal scale has only a minor role in the amount of uncertainty linked to catchment heterogeneity (Figures 3.5 and 3.8 to 3.10). Changing the time-step from 24- to 3-hour creates a larger variability, particularly for WaSiM simulations, which could be attributed to the more complex structure of WaSiM compared to Hydrotel. Contrary to the present time, the contribution of time-step in the variability of extreme flow in the future is considerable. This will be further explained in the next section.

6.2 Extreme flow projection

Based on Figures 4.3 to 4.4, global warming will reduce spring freshet peak flow (up to 60% decrease by the period 2081-2100, Figure 4.5). For summer-fall extreme flow, the results show an increase in magnitude in the future for small catchments, regardless of spatial resolution (Figures 4.6 and 4.7). As discussed in chapter 1, such a change is attributed to the increasing water-holding capacity of the atmosphere due to atmospheric warming and available humidity because of increased actual evapotranspiration (Yin *et al.*, 2018; Westra *et al.*, 2014).

Based on Figures 4.6 and 4.7, for subdaily simulations, there is a pattern regarding catchment size and the magnitude of change of extreme summer-fall flow. By decreasing catchment size, the magnitude of the extreme summer-fall flow increases by around 50%, and for large catchments, it decreases. The reason can be attributed to the shorter residence time of water in small catchments, and the spatial distribution of convective events that cover the whole area of such catchments at once . For larger catchments, convective rainfall is hardly relevant (Blöschl, 2022b). Given the average rise of AET (30% increase, Figure 4.5), the catchment wetness is already at the lower level leading to lower magnitudes of flooding in the future for large catchments. While this is true that the same conditions (i.e., higher AET) exist for small catchments, the flood mechanism for short-duration rainfalls is infiltration excess (as apposed to saturation excess for large catchments), for which soil wetness before intense rainfall is not relevant (Blöschl, 2022a).

6.3 Multi-scale regionalization

Chapter 5 suggests that the knowledge from gauged catchments can be transferred to ungauged catchments using a random forest regionalization model. In this context, using refined temporal resolution (3-hour time-step) showed a significant improvement in regionalization skill compared to daily time-step (Figure 5.6). In addition, the information can be transferred across spatial scales, i.e., from low resolution to high resolution. As a result, a parameter set can be obtained per HRU, which in turn enables launching simulations in a fully distributed mode. The results

show a benefit to using more spatially representative catchment descriptors for daily simulations (Figure 5.10).

Different groups of predictors are used (including meteorological, land cover, soil type, and topographic predictors) to approximate distributed hydrological model parameters. Comparing relative importance of these groups, no specific pattern emerged (Figure 5.7). However, the weights for each group are different. Soil type predictors have a lower importance compared to other groups. This is not in line with the results of Merz *et al.* (2020), in which no significant advantage over predictor groups was observed. The reason could be two-fold. Either the soil type of the study area is not diverse enough to influence catchment response, or the hydrological model is not sensitive enough to the soil type. Figure 6.1 shows the distribution of all predictors, which are separated by vertical black lines based on their own groups. In the column related to the soil type predictors, narrow distributions with outlier values are detectable for silt (SL) and clay (CL), showing a lack of diversity regarding soil type. Meteorological, topographic, and land cover predictors have wider distributions than soil predictors, which have greater powers to approximate hydrological model parameters based on Figure 5.7.

6.4   Recommendations for future studies

The limits of this research are as follows: first, the hypotheses are examined over a limited number of catchments located in Quebec, Canada. Therefore, the results cannot be extrapolated to locations with different climate conditions or flood regimes. Second, due to the considerable computational costs of calibration and simulation of the process-based hydrological models, the number of catchments studied is inevitably limited. The results are valid for these case studies, and adding other catchments may change these results. Third, we used one large-ensemble regional climate model for flood projections. Given that climate models are a major source of uncertainty in the projections, adding more models may change the projection results. Fourth, the uncertainty regarding observational data and using different datasets for calibration and bias correction were not considered.
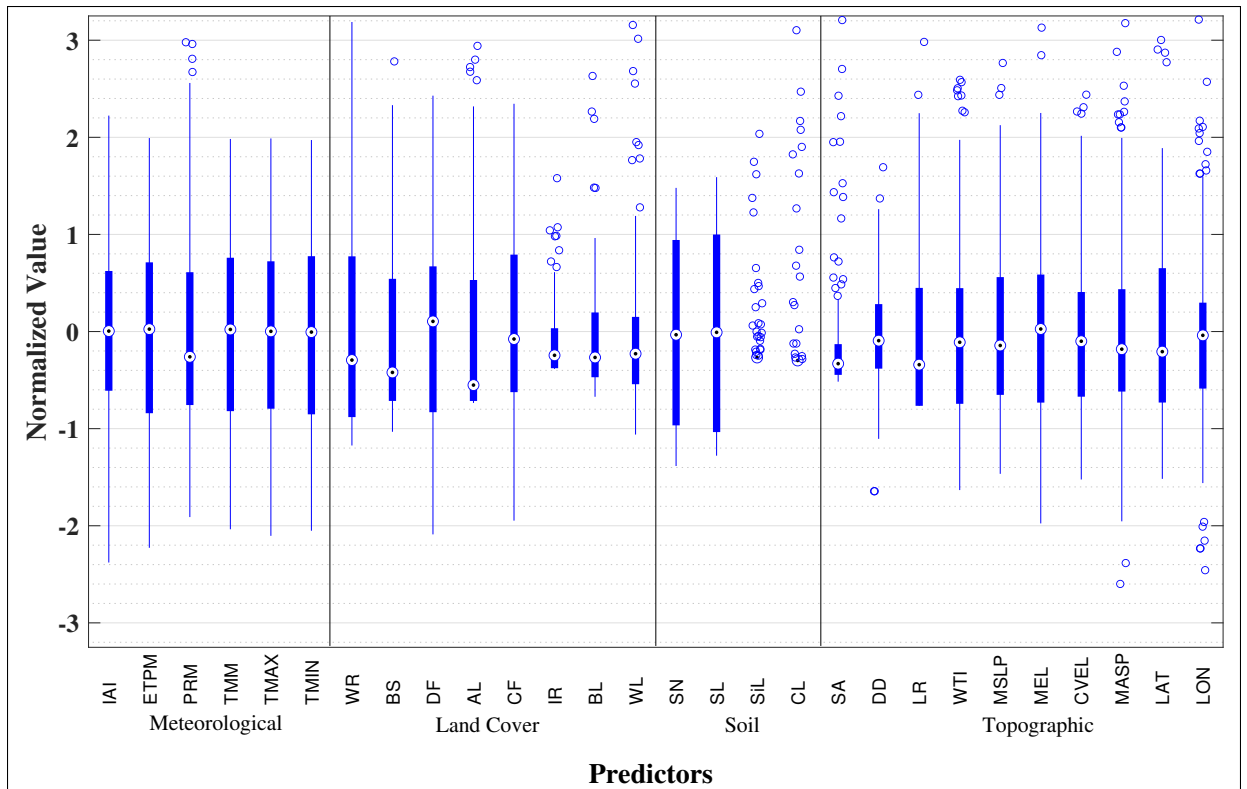
Figure 6.1   Distribution of catchment descriptors (predictors)
used for regionalization

The main issue in dealing with the application of process-based distributed hydrological models is their computational costs. While dramatic progress regarding computational power has been achieved in the last few decades, using these models still seems a daunting task. For example, calibration of a single catchment with a complex model like WaSiM may take weeks, if not months, even if high speed CPUs are used. Surrogate models can be a solution to this problem. Multiple approaches exist to create surrogate models for reducing the modeling cost. One approach is to use "low-fidelity" models to represent the original or "high-fidelity" model. The "high-fidelity" is the most accurate and desirable option of the modeling, and the "low-fidelity" model is a cheaper alternative to that option with adequate accuracy (Razavi, Tolson & Burn, 2012; Huot, 2019).

In this study, we proposed a method to create an ensemble of simulations for process-based models. The method consists in using a combination of different parameter resolutions (obtained from calibration) and catchment descriptors resolutions. The results show that parameters are transferable across scales for simulations. In this sense, the parameters calibrated from a low-resolution scale, i.e., the "low-fidelity" model, can be used for a high-resolution simulation, i.e., the "high-fidelity" model. This resulted in a sharp decrease in the computational time of the calibration. Figure 6.2 shows how coarsening spatial scale can significantly reduce the run-time for an 8-year simulation of the Aux Brochets catchment (a medium catchment) with WaSiM. In this figure, R with a number represent the spatial resolution in $m2$ used for calibration, with a 3-hour (3h in the figure) or 24-hour (24h) time step. Note that this approach can be recommended for large catchments ($> 500 km^2$) as it may cause distortion in simulations for small catchments according to Figures 3.8 to 3.10.

Investigating the impact of spatio-temporal scale in conjunction with the structural uncertainty is a potential research topic. This research suggests that the choice of hydrological model is essential in determining the variability of simulations with different spatio-temporal resolution. While multiple papers have investigated structural uncertainty (e.g. Seiller *et al.*, 2017; Chlumsky, Mai, Craig & Tolson, 2021; Gupta & Govindaraju, 2019), they have not focused on spatio-temporal issues, for instance how increasing the model's complexity interacts with the uncertainty linked to the catchment heterogeneity.

Another opportunity for future works is using Convection Permitting Regional Climate Models (CPRCM) to overcome the shortcomings of the current RCMs in the projection of sub-daily rainfall. CPRCMs are high-resolution climate models (<4km spatial resolution and sub-hourly temporal scale), enabling the representation of deep convective process, which is a significant cause of uncertainty in climate modeling (Ban, Schmidli & Schär, 2014; Foley, 2010; Ban *et al.*, 2021). The added value of using CPRCMs in impact studies for the research requiring high resolution, (e.g. flash flood projection or impact assessment of urban catchments) has been demonstrated (Lucas-Picher *et al.*, 2021). Moreover, models' high spatial and temporal resolutions provide the opportunity to investigate further the impact of spatio-temporal resolution
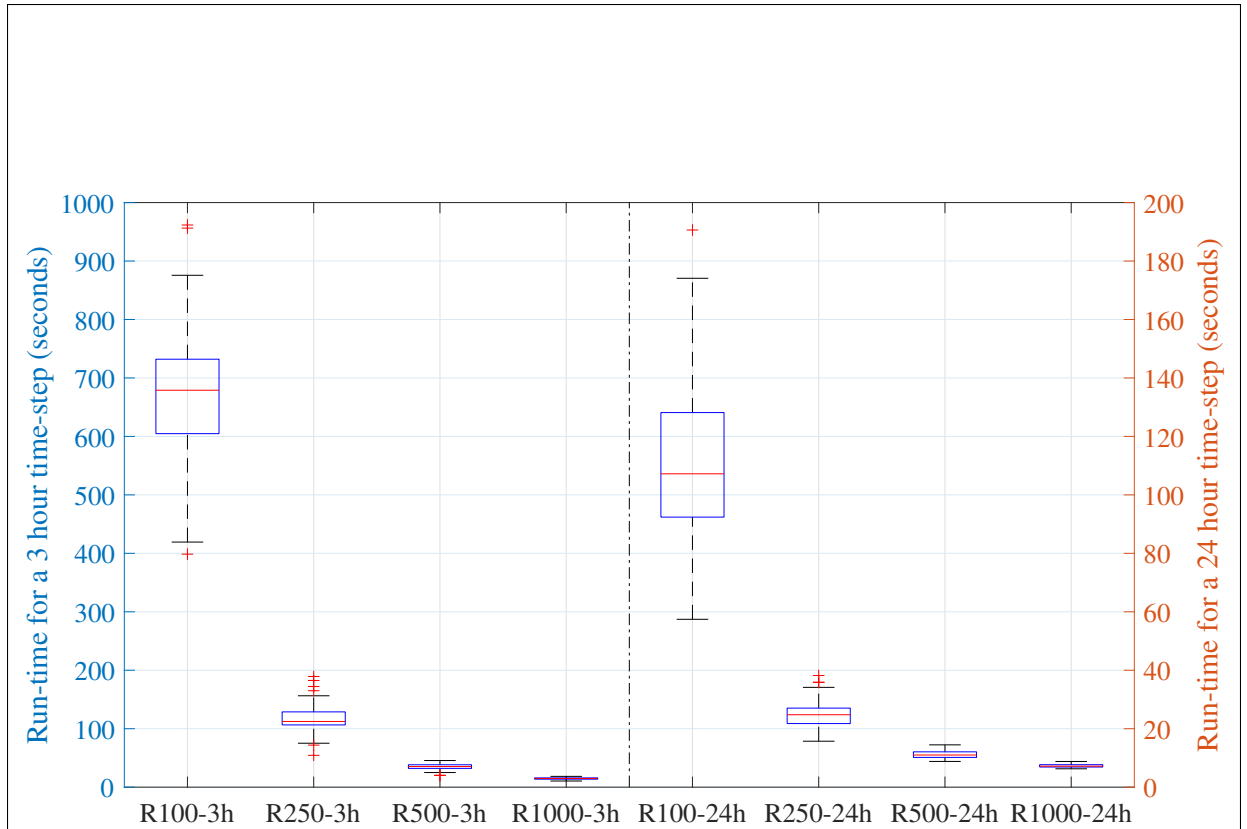
Figure 6.2   Comparing the run-time of simulations with
WaSiM at different spatial scales for the Aux Brochets
catchment

regarding the uncertainty linked to catchment heterogeneity. Since a large storage capacity is required to store the dataset, the use of such models is still limited to local and decadal projections. It is predicted that this restriction will be relieved during the next decade, which opens the avenue for multi-decadal and regional impact studies (Gutowski *et al.*, 2020; Schär *et al.*, 2020).

Finally, the work of Kratzert *et al.* (2018, 2019b) on Long Short-Term Memory (LSTM) neural networks opens an exciting avenue for the application of deep learning in hydrology. Recent studies have shown that LSTM neural networks can outperform traditional conceptual-based and process-based hydrological models in terms of streamflow forecasting (Arsenault *et al.*, 2022; Mai *et al.*, 2022a). In addition, LSTM shows minimal loss of efficiency when applied to

ungauged catchments (Arsenault *et al.*, 2022). In general, the application of deep learning and specifically LSTM can significantly improve the shortcomings of simulation and forecasting in ungauged basins (Arsenault *et al.*, 2022). To address the black box nature of statistical learning methods, the 5th chapter of this thesis successfully tested a combination of RF and a process-based hydrological model. Such approach can be used for other applications of LSTM in hydrological modeling, given a large dataset is available. As such, the underlying physical processes that cannot be captured by the hydrological models, can be approximated by pure numerical models such as machine learning and deep learning.

CONCLUSION AND RECOMMENDATIONS

The 2021-2022 IPCC report accumulated extensive evidence corresponding with the impacts of climate change around the world (Masson-Delmotte *et al.*, 2021). It is estimated that flooding patterns will be altered for snow-dominated catchments. This necessitates accurate flood simulation to devise coherent adaptation and mitigation policies. Hydrological models have gained more importance in climate change impact assessment studies as the final link of the hydro-climate modeling chain. With the continuous progress in computational power, the application of process-based distributed hydrological models has dramatically increased over the recent years. However, scale issues, which can be summarized in nonlinear catchment responses at different spatial and temporal scales, make it challenging to provide an accurate representation of physical processes in hydrological models. The present research aimed at quantifying the impacts of the spatio-temporal scale of modeling in streamflow simulation, flood projection, and regionalization using process-based distributed hydrological models. In the next paragraphs, research hypotheses are restated (in bold text) followed by the related conclusions. Finally, practical recommendations are proposed.

1. **Hypothesis 1: The choice of a level of spatio-temporal discretization alters model parameters, which leads to uncertainty in streamflow and flood simulation. By increasing catchment area, the contribution of the choice of spatial scale and hydrological model in such uncertainty increases, and that of time-step decreases.** Chapter 3 proposed a new approach to create an ensemble of simulations by changing the spatio-temporal resolution of the model. The results show that there exists uncertainty corresponding with the heterogeneity of the catchments in the simulation of low, high, and extreme flow in the historical period. The study concludes that (1) there is no pattern regarding the catchment size and the uncertainty linked to heterogeneity. (2) Refining the time step of the modeling has only a minor impact on increasing the variance of the simulation ensemble in historical time period. (3) The more sophisticated model in terms of representation of

hydrology processes demonstrates more sensitivity to the spatio-temporal scale. Chapter 4 confirms that (1) there is no evidence of a relationship between the catchment size and spatio-temporal scale in the variability of flood projections (i.e. in the future time periods). (2) By increasing catchment size, the contribution of the hydrological model to the variance of simulations increases. Overall, the first part of the hypothesis regarding the existence of uncertainty linked to the spatio-temporal scale can be verified for this case study. But there is no evidence of the existence of a link between that uncertainty and catchment size.

2. **Hypothesis 2: For small catchments ($< 500 km^2$), refining temporal resolution of simulation (from daily to subdaily) increases the relative change (from reference to future) of extreme summer-fall flow. Refining temporal resolution will not significantly affect projected extreme summer-fall flow for large catchments ($> 1000 km^2$).** Chapter 4 confirms that, contrary to large catchments, by refining the time-step in modeling small catchments, the magnitude of extreme summer-fall flow in the future exceeds the reference period and the relative change is higher compared to daily simulations. Therefore, hypothesis 2 is verified for this case study.

3. **Hypothesis 3: There do exist relationships between model parameters and catchment descriptors that can be approximated by the random forest method. The underlying information related to catchment characteristics is transferable across scales through that approximation. Using fine-scale (in time and space) catchment descriptors improves the skill of the regionalization model.** Chapter 5 investigates the third hypothesis by building a random forest based regionalization method, which is a combination of a distributed model and machine learning. The method estimates the parameters of Hydrotel from the physiographical and meteorological characteristics of gauged catchments. The results support the third hypothesis, as the information regarding catchment characteristics captured in the model parameters in gauged catchments can be transferred to pseudo-ungauged catchments. Moreover, using a refined time scale for regionalization significantly

reduced the loss of efficiency between calibrated parameters and those obtained from regionalization. Additionally, using a nested catchment framework, the results suggest that the information can be transferred from low-resolution lumped catchment to high-resolution RHHUs. In this regard, if more spatially representative catchment descriptors are used (fully distributed model), the regionalization skill is improved for the 24-hour time-step but not the 3-hour time-step.

Overall, this study demonstrates the importance of representing spatio-temporal resolution in process-based hydrological models for accurate streamflow simulation, regionalization, and flood projection. This allows us to direct more attention to scale issues and focus on variation of hydrological variables over scale. This should be seen especially in light of emerging high-resolution climate models, in which their resolution is about to close (or surpass) the gap between spatial resolution of climate and gridded data. This effort, therefore, may result in a more realistic streamflow projection in the future and improve our ability for decision-making in the face of climate change. The main practical recommendations from this thesis are:

- Multi-scale simulation should be considered while using a sophisticated process-based model, especially for flat catchments or catchments with a large portion of farmlands. Given that the contribution of the model in the variance of simulations increases by the catchment size, multi-model simulation for large catchments is recommended.

- Refining time-step from daily to subdaily causes a significant increase in the flood projections for small catchments. Subdaily time-step should be considered for impact studies on small catchments.

- Random forest can be used as a powerful tool for regionalization. The method shows promise in regionalization when when the number of catchments are limited. Contrary to artificial neural network and deep learning, random forest is not entirely "black box" and the relationships between model parameters and catchment descriptors are interpretable.

This helps better identifying important catchment characteristics and remove those with less importance in approximating model parameters.

- This thesis suggests that refining the temporal scale significantly improves regionalization skills and should be considered in practice.

- Using refined catchment descriptors should be considered for regionalization of a distributed model in daily simulations. This might help improve the regionalization skill particularly for more diverse catchments in terms of topography and land use.

Figure I-1 shows the Hydrotel2 annual hydrograph simulations when the number of HRUs varies. In general, a slight widening of the uncertainty bounds can be observed, manifesting a higher sensitivity of the Hydrotel2 set-up to changes in spatial resolution as compared to the Hydrotel1 simulations. Table I-1 shows the statistics of model efficiencies. KGE values are
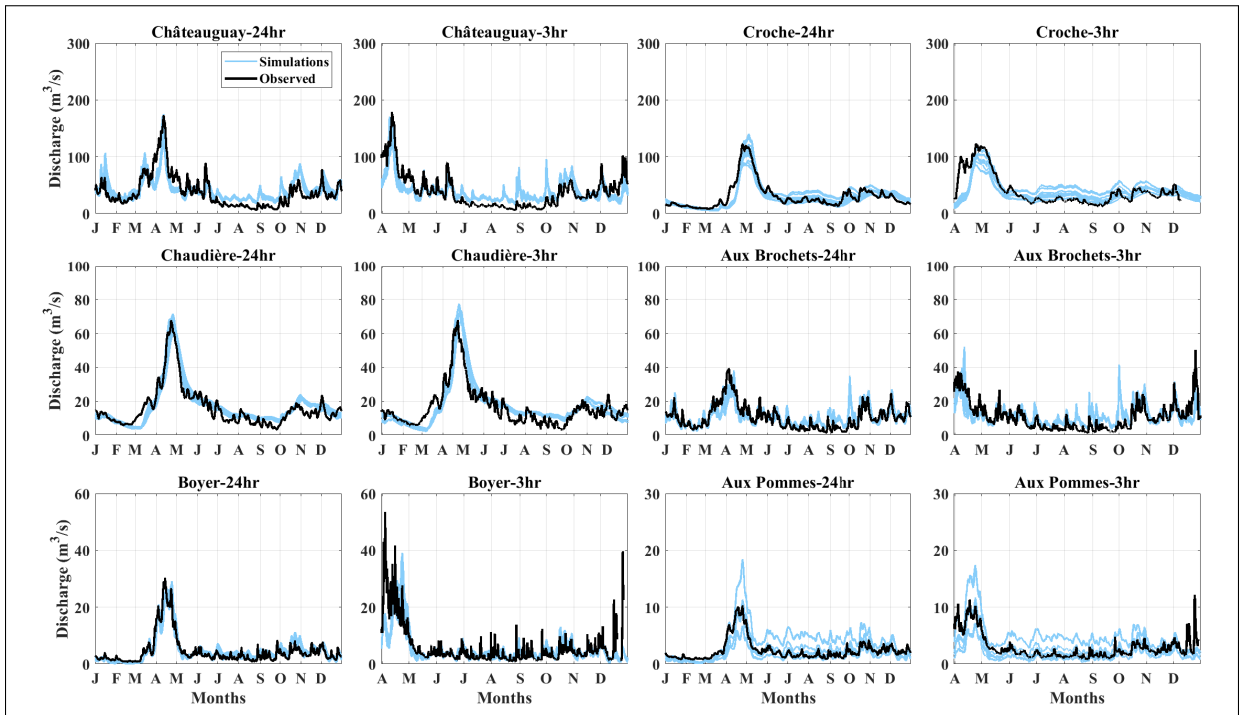
Figure-A I-1   Annual hydrographs of the selected catchments simulated by Hydrotel (Hydrotel2) and compared to observed data. The modeling time-steps are 24 and 3 hours. The responses are arranged according to the size of the catchments: large catchments (>1000 km^2) are on the top row; medium-sized catchments (between 500 and 1000 km^2) are on the middle row; small catchments (<500 km^2) are on the bottom row

consistent between validation and calibration periods for all model-catchment pairs for both 3- and 24-hour time-steps. For the majority of cases, the STD of the KGE for 3-hour is higher than for the 24-hour time-step. The Wilcoxon rank sum test was used to investigate the statistical significance of this change in KGE, with the null hypothesis being that the two ensembles are

equal at the default 5% significance level. The bold numbers in the last two columns of Table I-1 indicate that the change between the two distributions is significant (i.e. p-value<0.05). As can be seen, a few catchments show a significant change in the efficiency between the 24- and 3-hour time-step. Moreover, comparing the STD of models, no significant differences between them can be observed.

Table-A I-1   Land use and topographic characteristics of the catchments

Catchment Characteristics	Boyer	Chaudière	Aux Brochets	Châteauguay	Croche	Aux Pommes
water (% of surface area)	0.01	0.05	0.02	0.01	0.05	0.01
bare soil (% of surface area)	0.01	0.02	0.00	0.01	0.06	0.02
deciduous forest (% of surface area)	0.14	0.48	0.34	0.21	0.34	0.29
agriculture lands (% of surface area)	0.66	0.04	0.45	0.47	0.00	0.31
coniferous forest (% of surface area)	0.08	0.33	0.13	0.03	0.46	0.25
impermeable (% of surface area)	0.03	0.02	0.03	0.03	0.01	0.06
bog land (% of surface area)	0.06	0.02	0.00	0.01	0.04	0.04
wetland (% of surface area)	0.01	0.04	0.03	0.03	0.03	0.02
sand (% of surface area)	0.56	0.04	0.50	0.01	0.75	1.00
sandy loam (% of surface area)	0.44	0.96	0.45	0.61	0.25	0.00
silt loam (% of surface area)	0.00	0.00	0.00	0.08	0.00	0.00
clay loam (% of surface area)	0.00	0.00	0.05	0.30	0.00	0.00
drainage density [km/km^2]	0.38	0.38	0.42	0.32	0.33	0.27
lake ratio[km/km^2]	0.00	0.05	0.01	0.00	0.01	0.00
wetness index	9.96	8.53	9.26	10.38	8.03	9.60
mean slope	1.59	7.68	3.65	1.13	7.30	2.46
mean elevation	119.52	541.71	134.01	76.11	385.20	133.60
coefficient of variation of elevation	0.41	0.24	0.54	0.91	0.18	0.44
median aspect	275.00	173.00	231.00	179.00	177.00	188.00
latitude [decimal degree]	46.75	45.45	45.05	45.15	48.10	46.75
longitude [decimal degree]	-70.95	-70.90	-72.90	-73.95	-72.60	-71.60

We used the empirical orthogonal function (EOF) for decomposing the signal of AET in time and space for WaSiM and Hydrotel1 at Châteauguay (24-hour time-step). The results are shown in Figures I-2 and I-3. Figure I-2 shows the dominant spatio-temporal pattern of AET for different spatial resolutions at the monthly time scale (2001-2017), according to a principal component analysis (PCA). In this figure, the left column represents the leading spatial distribution characteristic of EOF, and the right column shows the temporal variation of the principal component (PC). For all resolutions, the explained variance by the leading component is larger than 98%. Therefore, we ignored the impact of other components with small contributions in the mode of variability. Distributions of EOF across the catchment at different scales show a large difference in the magnitude of simulations with different scales.

For simulations with a resolution of 100m2, for example, the magnitude of the first mode of AET is near zero and rather uniform across the catchment. By increasing the scale to 1000 m, the magnitude of the EOF increases (in terms of absolute value) and more variability across the catchment can be seen. On the other hand, a high degree of oscillation in time is visible for the 100 m resolution compared to 1000 m.
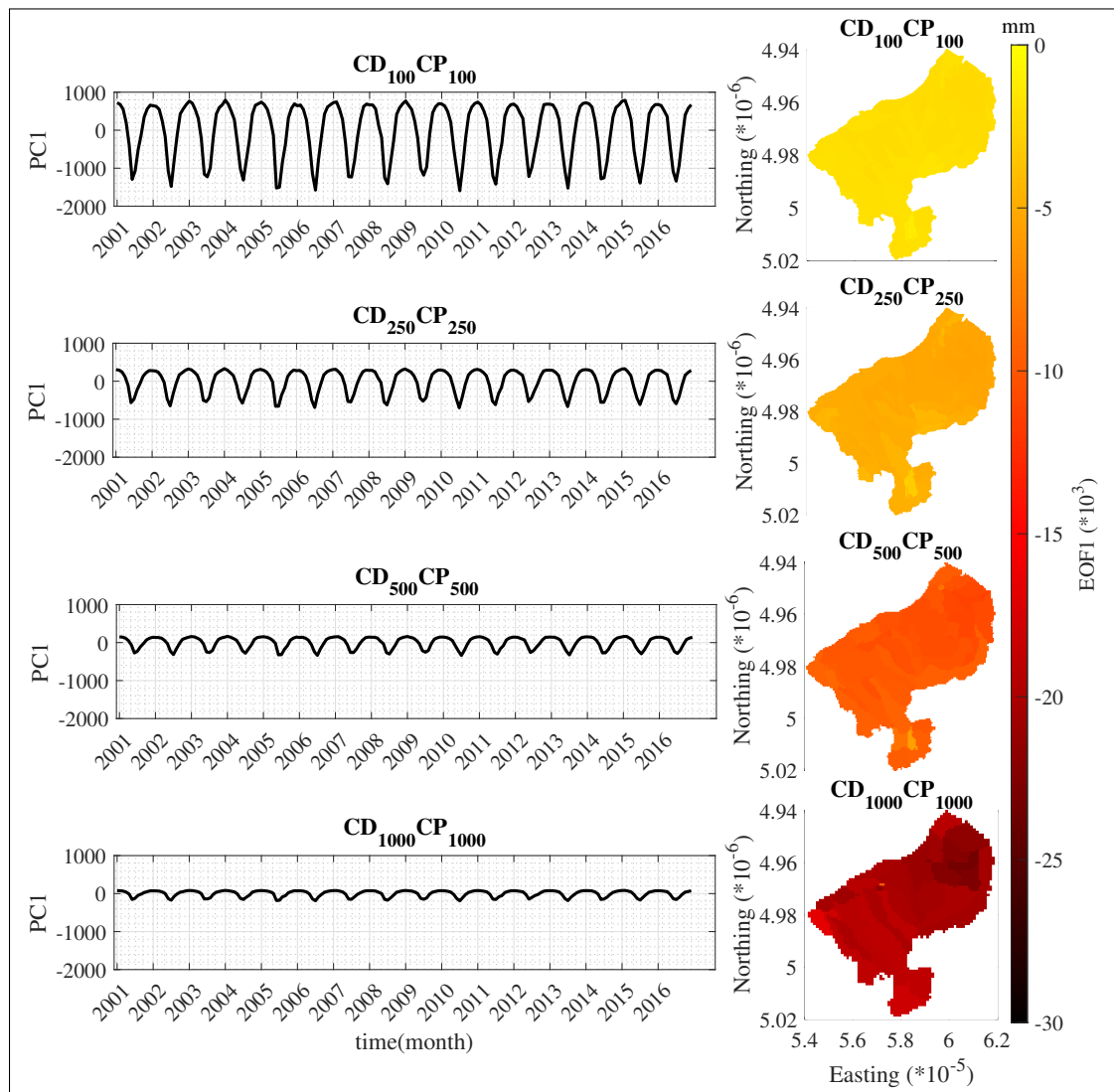
Figure-A I-2   The first mode of empirical orthogonal function (EOF) and corresponding principal component (PC1) derived for the Châteauguay catchment simulated by WaSiM

Figure I-2 shows EOF and PC components of AET simulated by Hydrotel1 at Châteauguay (24-hour time-step). The experienced variance of the leading EOF is more than 96%. Similar to WaSiM, a large difference in terms of the magnitude of EOF can be seen across scales. Similarly, the magnitude of simulations with 100 *m* is near zero, and by increasing the scale, the absolute value of EOF increases. Regardless of the magnitude of the EOF, a similar pattern in the spatial distribution of EOF can be seen across scales. The pattern is more visible comparing the simulations 250, 500, and 1000 *m* resolutions. This pattern resembles the spatial pattern of AET in Figure 3.7. Regarding PCs, the behavior of Hydrotel1 simulations is similar to WaSiM: a high amplitude of oscillation when the resolution is the high and low amplitude of oscillation by decreasing the resolution. However, the temporal variability is lower than for WaSiM in terms of amplitude. Moreover, the PCs of Hydrotel1 show a higher frequency compared to WaSiM. Figures I-4 and I-5 illustrates a separation of the total uncertainty for extreme streamflows into contributions from CDs and CPs using Hydrotel1 and Hydrotel2, respectively. Figure I-4 illustrates the decomposition of uncertainty for extreme streamflows simulated by Hydrotel1. The magnitude of MDE for both CDs and CPs is limited compared to WaSiM. Likewise, the MDE of CPs is larger in most cases, except for the Aux Brochets catchment with a 3-hour time step and the Aux Pommes catchment. In general, Hydrotel2 (Figure I-5) simulations show larger MDEs than Hydrotel1 (Figure I-5) simulations. Also, the number of cases in which the dominant source of uncertainty is CDs is increased (compared to WaSiM) as the Châteauguay and Aux Brochets catchments show larger MDEs across the vertical axis (note that the MDEs of CDs calculated for QT50 for Aux Brochets-3 hour are larger than 1, and have been removed for the sake of consistency in comparisons). As Figure I-5 shows, the range of uncertainties corresponding to these two catchments is substantially larger than for other catchments, in which the dominant source of uncertainty is CPs.

As discussed in Figures 3.14, I-4 to I-5, the dominant cause of uncertainty in the simulation of extreme streamflow relates to CPs resolution for most of the catchments. There are exceptions in which the dominant source of uncertainty in the simulation of those extreme values can be attributed to changes in the resolution of CDs. Catchments such as Aux Brochets, Aux Pommes,
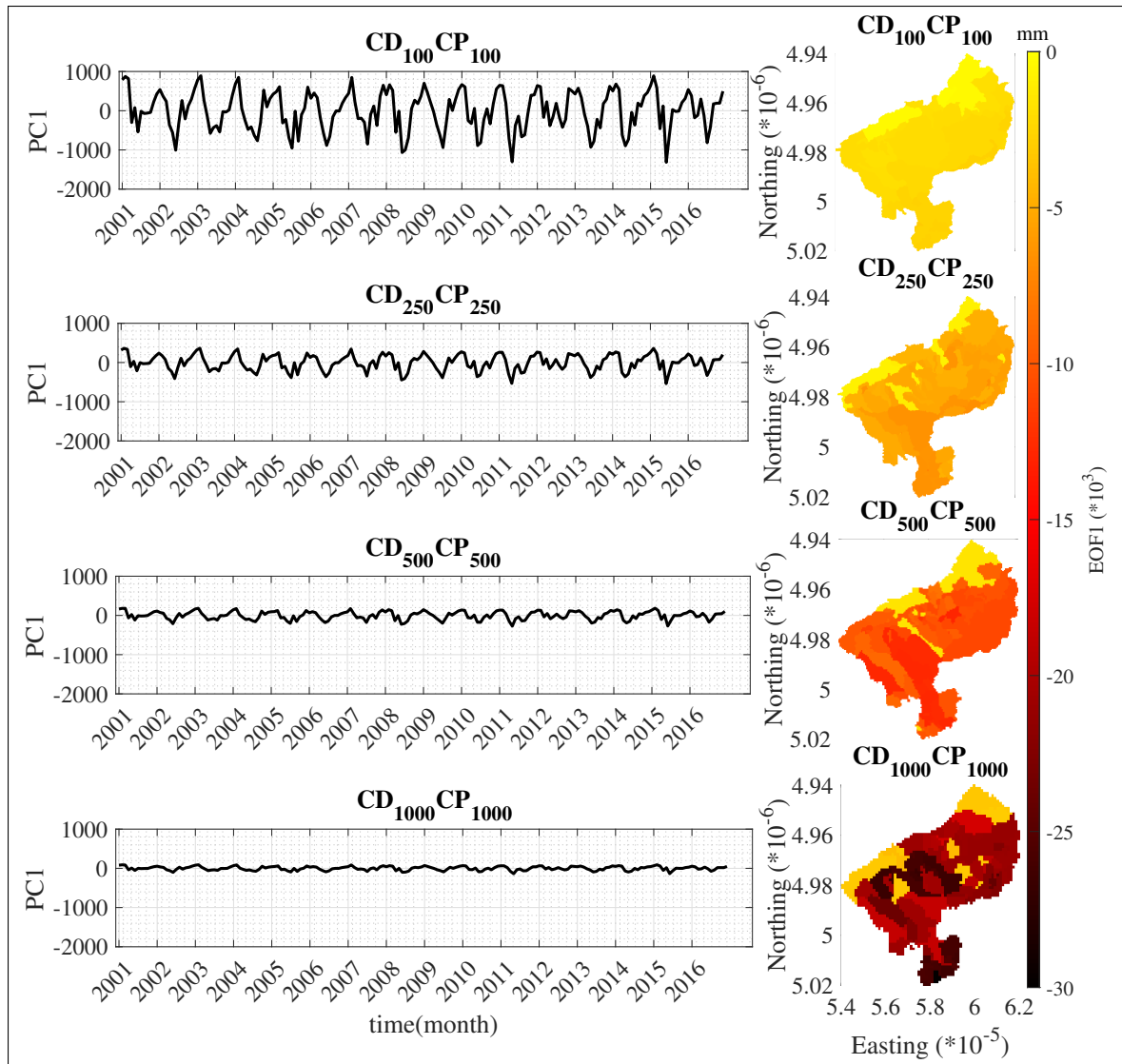
Figure-A I-3   The first mode of empirical orthogonal function (EOF) and corresponding principal component (PC1) derived for the Châteauguay catchment simulated by Hydrtel1

and Châteauguay, are among these cases. From this list, the Aux Brochets (3-hour) catchment demonstrates the highest level of dominance of CDs, regardless of the model or configuration used for simulations. To explore the reason for the observed sensitivity, we used simulations from WaSiM, as this model offers further insights regarding the changes in state variables and fluxes across the catchment.
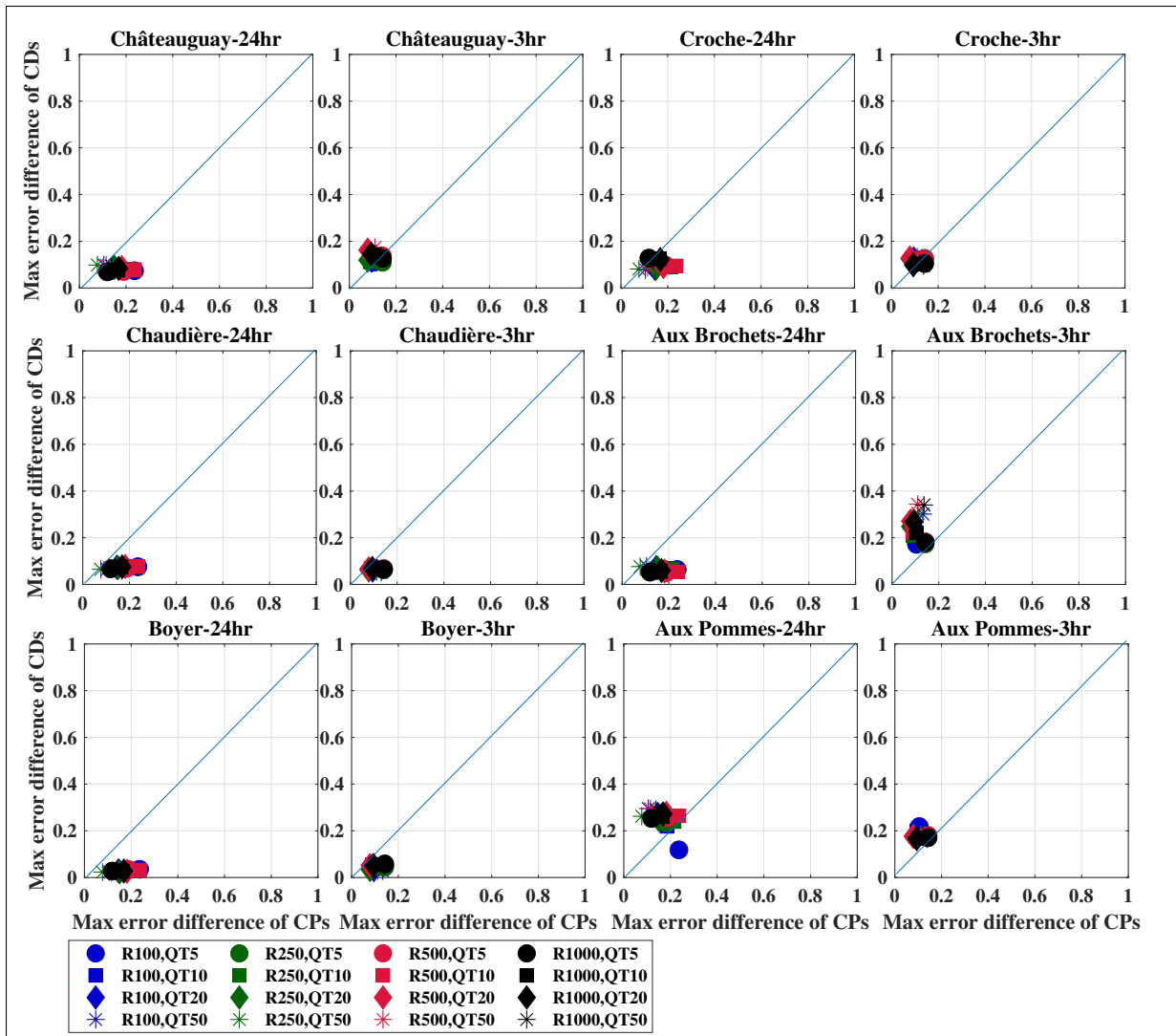
Figure-A I-4 Relative errors for the simulation of summer-fall floods with 5-, 10-, 20-, and 50-year return periods using the Hydrotel1 configuration. QT represents a flood with a specific return period. For instance, QT5 is the flood magnitude corresponding to a 5-year return period. R represents the resolution of CDs or CPs, in which the Maximum Error Difference (MED) is calculated

Figure I-6 shows the distribution of average groundwater levels across the catchment. In each column, the resolution of CPs is constant, while the resolution of CDs is changing. By coarsening the resolution, a major increase in groundwater level near the outlet of the catchment (located in the north-western part) can be observed. For instance, the distribution of groundwater across the catchment in subplots a and e is similar, and it changes for subplots i and m. This change in the
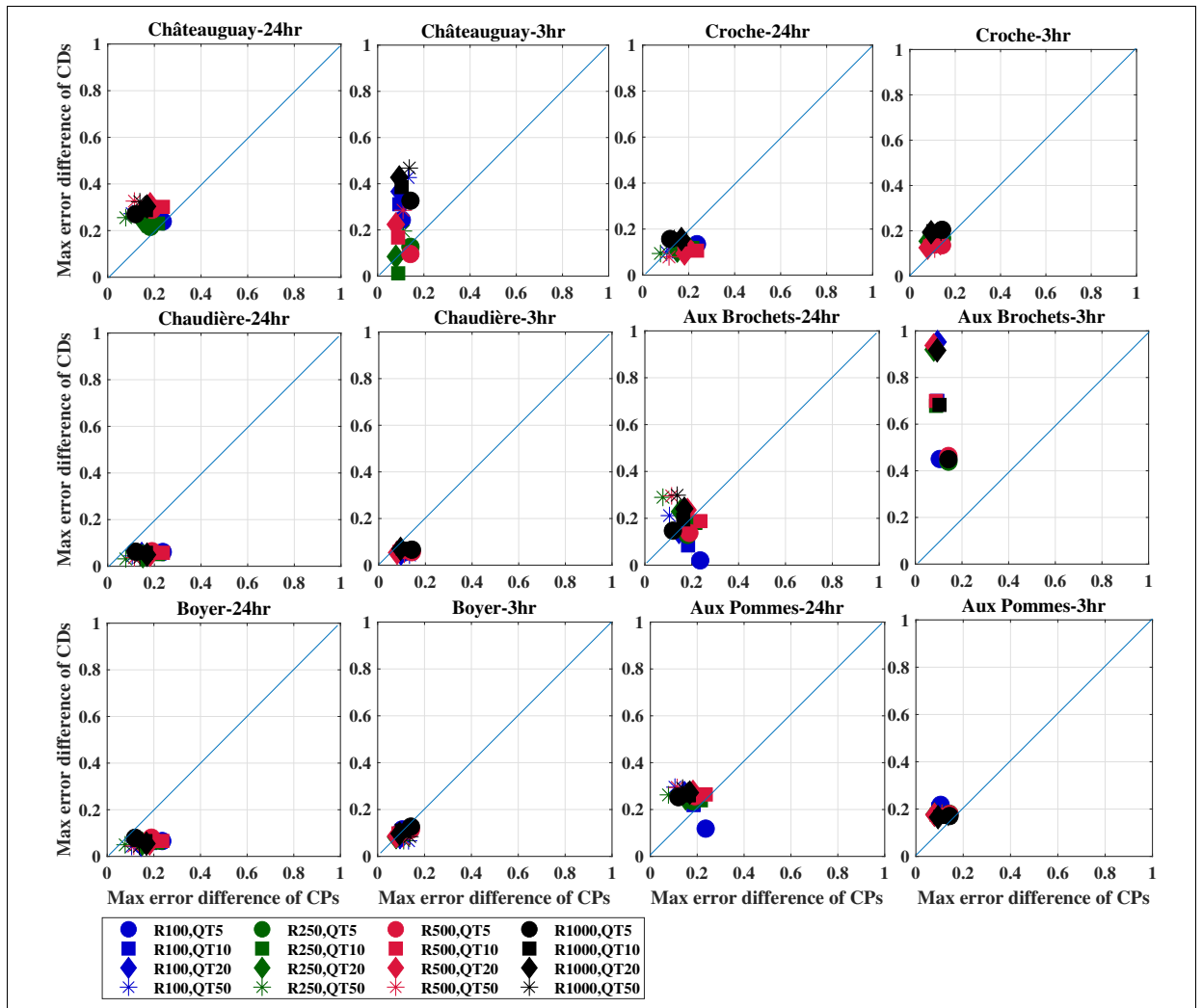
Figure-A I-5   Relative errors for the simulation of summer-fall floods with 5-, 10-, 20-, and 50-year return periods using the Hydrotel2 Configuration. QT represents a flood with a specific return period. For instance, QT5 is the flood magnitude corresponding to a 5-year return period. R represents the resolution of CDs or CPs, in which the Maximum Error Difference (MED) is calculated

distribution of groundwater across the catchment can also be seen for other CP resolutions (e.g., b, f, j, n, etc.).

To explore further, we selected the groundwater distribution results for 100 and 500 *m* CDs as representative of high and low-resolution catchment descriptors. We compared them with the distribution of slopes across the catchment. Figure I-7 shows the average groundwater level
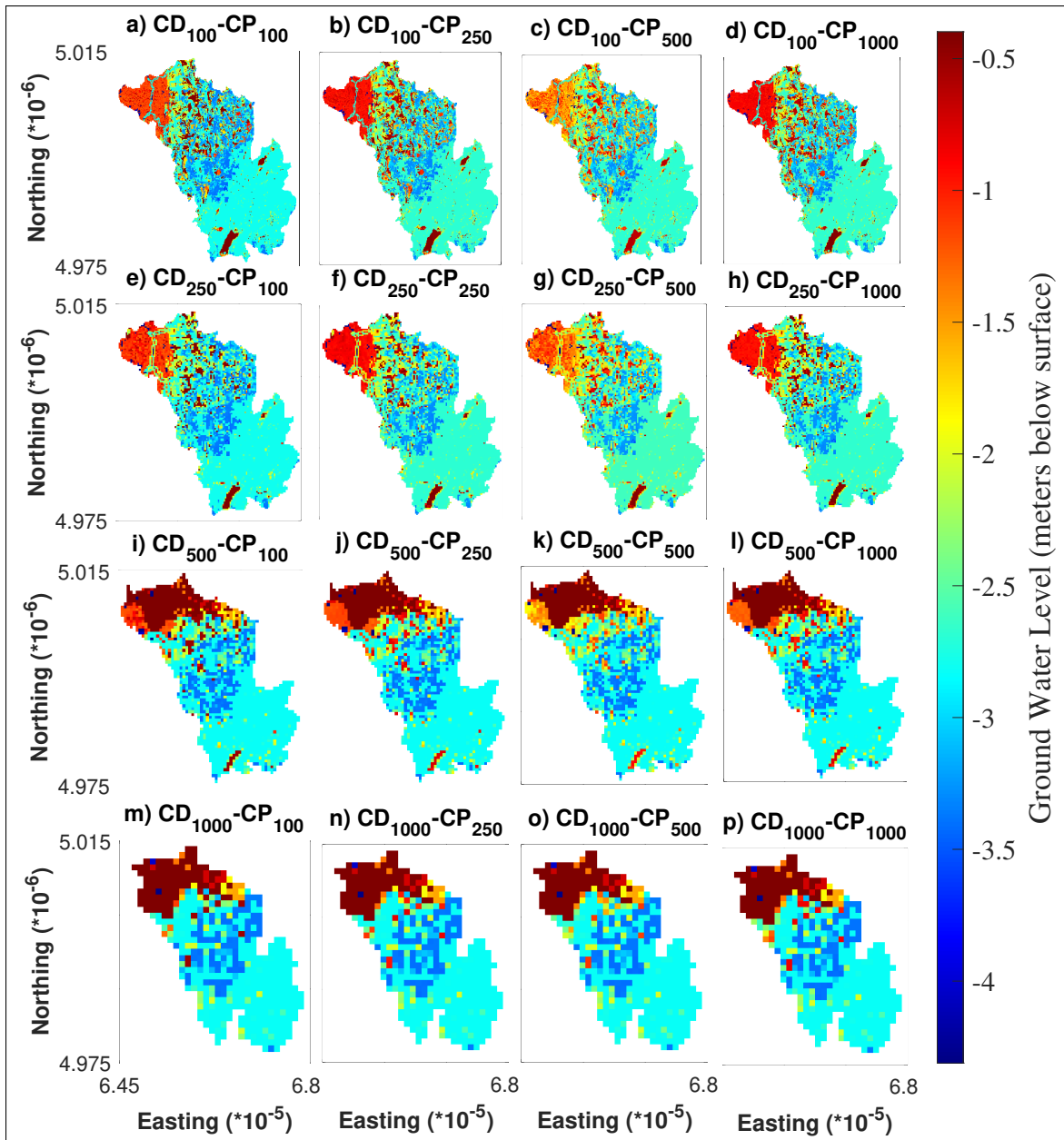
Figure-A I-6   Distribution of groundwater elevation across the Aux Brochets
catchment for different resolutions for a 3-hour time step

(bottom row) and slope (top row) within the catchment. Subplot c (CD 100 *m*) shows that the

maximum groundwater level can be found in the middle part of the catchment. Nevertheless,

for subplot d (CD 500 *m*), most of the groundwater accumulates in the downstream part of the

catchment. This can be explained by looking at the top row showing the slope distribution. In

subplot a (100 *m* resolution), there are small-scale hillslopes and valleys, which spatially correlate with the maximum groundwater level in the middle of the catchment. These uneven areas that retain groundwater at specific parts of the catchment disappeared during the interpolation for 500 *m* CDs (subplot b), resulting in an accumulation of groundwater downstream.
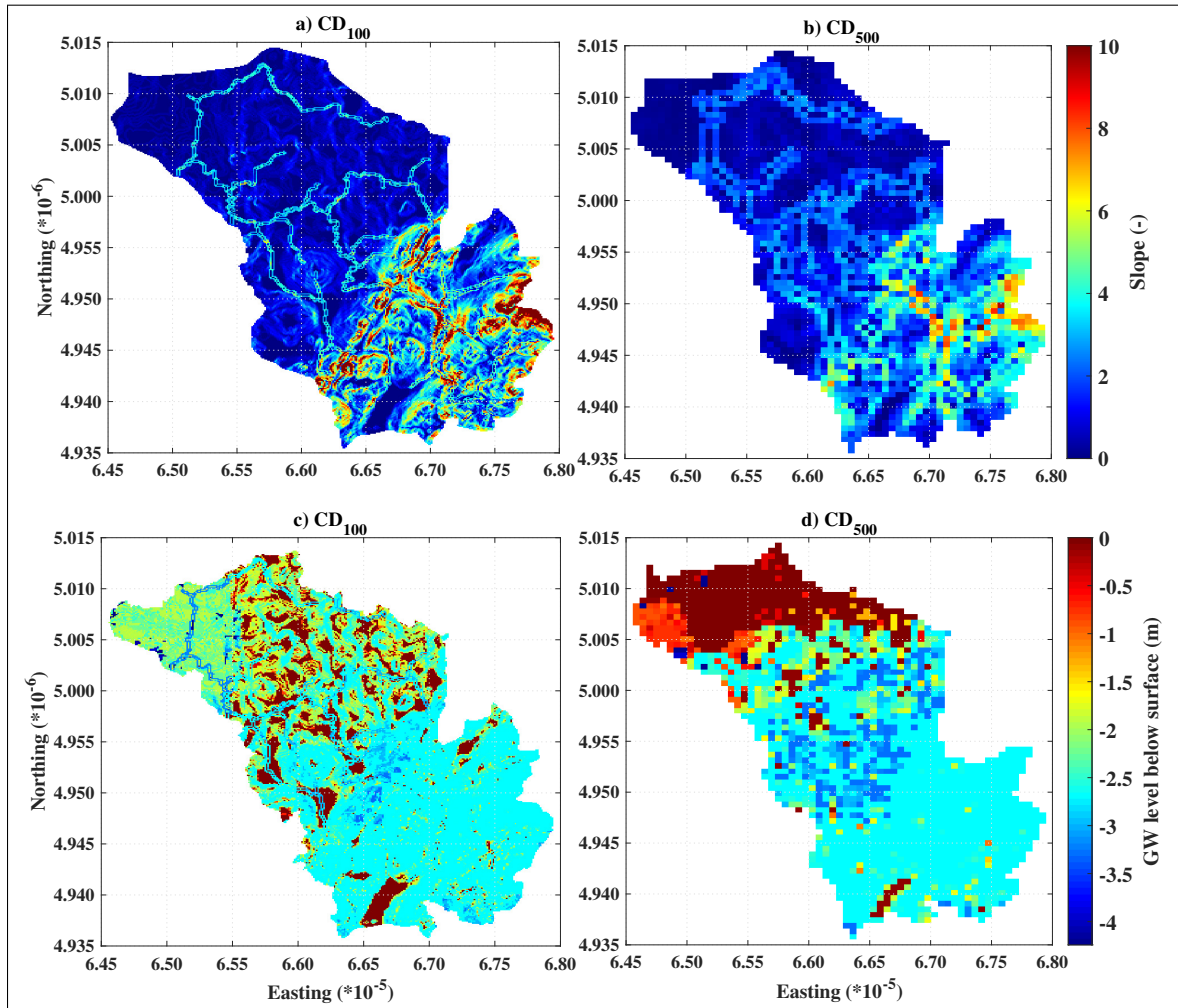
Figure-A I-7   Comparison of slopes (top) and groundwater elevation (bottom) for the Aux Brochets catchment (3 hr) simulated by WaSiM

Figure I-8 illustrates the catchment response at the outlet and at Reach1 (or R1 subbasin) for the spring freshet of 2008. R1 is located right before the outlet in the downstream area. Here, the dashed lines represent direct runoff from the subbasins, and the solid lines show the simulated streamflow at the 3-hour time step. In both subbasins, the catchment responses reproduced by

the 500m2 resolution demonstrate considerable fluctuations, particularly for the R1 subbasin. The reason for this is that the water table is very close to the surface in this area, and this reduces the damping effect of interflow and baseflow down to near zero. As a result, any change in the meteorological forcing translates into a direct flow and a corresponding rapid reaction of the catchment in the R1 subbasin. The fluctuations further transfer and commensurately affect the streamflow at the outlet of the catchment. In fact, changes in the resolution of CDs alter the magnitude and timing of the peak flow, regardless of the variations of CPs. Such behavior can explain the dominance of CDs over CPs in Figures 3.13, I-5, and I-6 for the Aux Brochets (3hr) catchment.
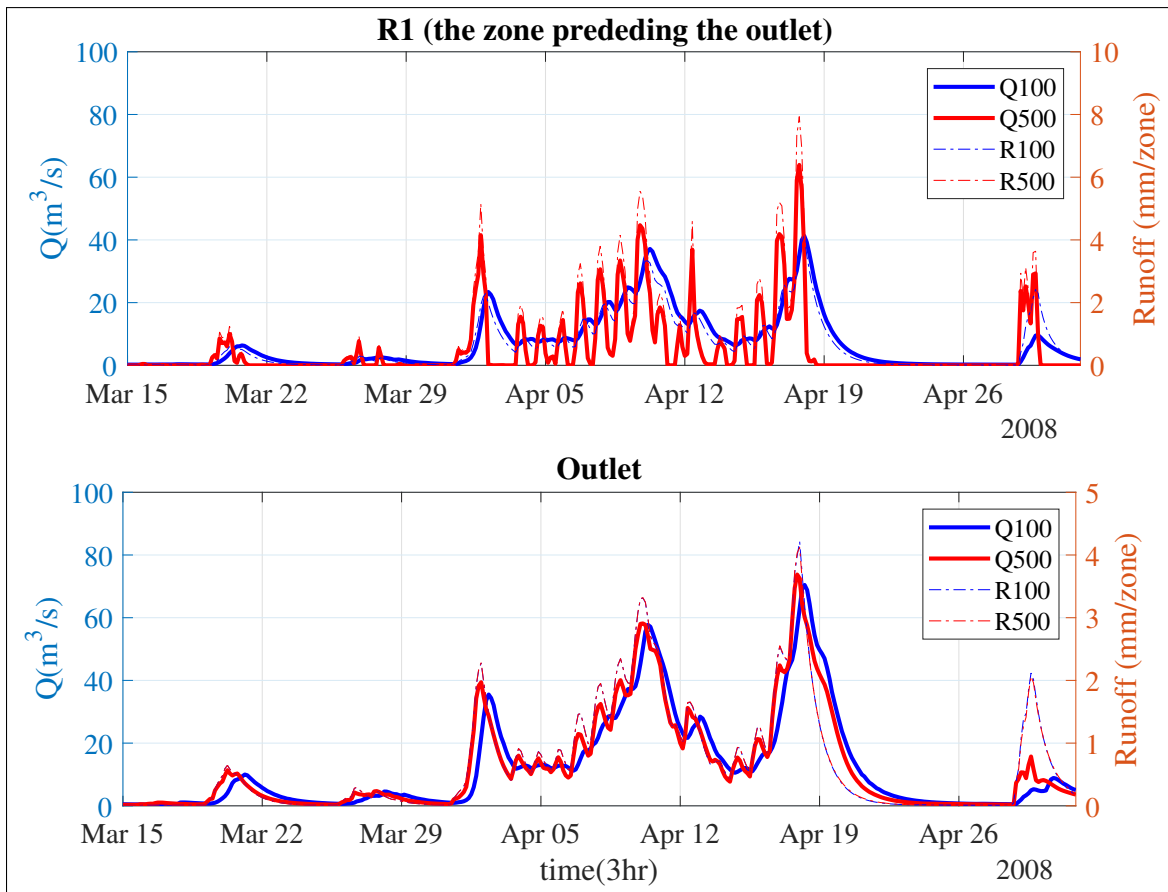
Figure-A I-8   Routed discharge (Q100, Q500) and direct runoff (R100, R500) of 100 and 500 m2 CD resolutions simulated by WaSiM for the outlet and Reach 1 (R1 is the reach located in the downstream area next to the outlet of the catchment) for the Aux Brochets catchment for a 3-hour time step

As discussed in Figures 3.8-3.10, certain poor simulations of Hydrotel for the Aux Pommes (3- and 24-hours) catchment are probably due to the incompatibility between the resolutions of CDs and CPs. Here we further analyze the catchment response to explore the reason for such incompatibility. Figure I-9, subplots a and b demonstrate the relative change of spatially averaged hydrological variables to the simulation CP100CD100. Subplot a shows the change in all combinations of CP100 (i.e. CP100CD250, CP100CD500, CP100CD1000) and subplot b shows the change in all the combinations of CD100 (i.e. CD100CP250, CD100CP500, CD100CP1000). Subplot c of Figure I-9 shows the mean annual hydrograph for all combinations of simulations for a 24-hour time-step. In this subplot, the lines for the simulations with CP100 or CD100 are thicker and separated with blue and red colors, respectively.

Figure I-9 panel a shows that when CP100 is kept constant and CD varies, lateral flow undergoes a significant change (particularly for the case of CP100CD1000). Lateral flow is the flow on sub catchments (vertical simulation units) towards the river (Fortin *et al.*, 2001a), which is simulated using the kinematic wave equation. According to this equation A I-1, the lateral flow depends on the slope of the cell and on Manning's coefficient, both of which are sensitive to variations in land use and DEM.

$$h = [\frac{n}{1.49\sqrt{S_0}}]^{3/5}R^{3/5} \qquad\qquad (\text{A I-1})$$

In A I-1, R is lateral flow from cell to cell (m^2s^-1); h is flow depth (m); n is Manning's roughness coefficient; S0 is the slope of the cell.

By increasing the scale of CDs, the change of slope and \or Manning's coefficient causes a significant increase of lateral flow, resulting in an overestimation of streamflow in the spring and summer-fall periods according to subplot c (the blue lines).

Figure I-9 panel b shows that the average spatial distribution of variables is mostly influenced by ETPfor constant CD. In this case, the variation of the scale of the calibration parameter shows incompatibility with the scale of the CDs leading to an increase of ETP to 5 to 6 times higher
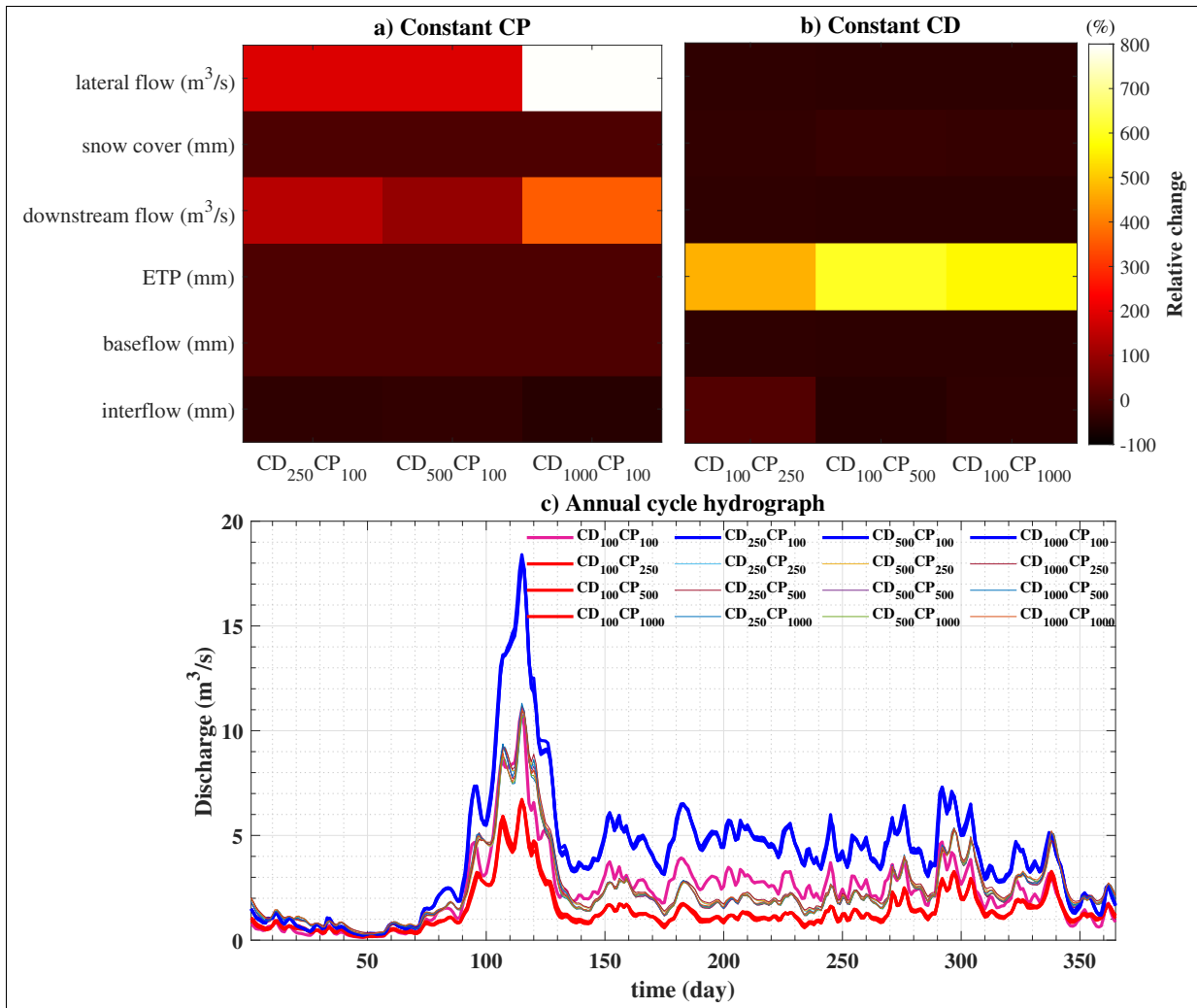
Figure-A I-9   Analyses of catchment Aux pommes by Hydrotel2 at a 24-hour time-step : a) spatially averaged hydrological variables for constant CP100 and variable CDs; b) spatially averaged hydrological variables for constant CD100 and varied CPs; c) annual cycle hydrograph at the outlet

than the original value for simulation CD100CP100. As a result, a decrease in the values of other hydrological variables (i.e., lateral flow, snow cover, base flow, and interflow) results in the underestimation of flow for the red lines.

Similar to the Boyer catchment simulated by WaSiM (Figure 3.15), Aux Pommes' characteristics in Table I-2 do not show distinct properties to help justify the catchment's unique response with regards to modeling options. Apart from having the smallest surface area, Aux Pommes has

the highest impermeability (6%) and the lowest drainage density (27 km/km^2). Additionally, this catchment has a uniform soil type (sand with low Manning's roughness). All these factors trigger the sensitivity of the catchment towards the change of CDs.

Table-A I-2   Table S1- Mean and Variance of KGE of all spatial simulations in 3- and 24-hour time-step for validation and calibration periods. a- The results under STD of KGE is a summation of KGEs' standard deviations in validation and calibration. b-The results of the Wilcoxon rank sum test to evaluate the change of ensemble of KGEs from 24- to 3- hour time-step for calibration and validation periods. Bold values show that the change is significant. *Inconsistent simulations were removed

models	catchments	Mean of KGE				STD of KGE[a]		Wilcoxon rank sum test[b]	
		cal-24h	val-24h	cal-3h	val-3h	24h	3h	p-value (cal)	p-value (val)
WaSiM	Châteauguay	0.8	0.76	0.81	0.81	0.04	0.02	0	0
	Croche	0.86	0.85	0.85	0.82	0.09	0.13	0.72	0.51
	Chaudière	0.82	0.77	0.82	0.77	0.03	0.04	0.19	0.64
	Aux Brochets	0.75	0.73	0.71	0.7	0.07	0.09	0	0.16
	Boyer*	0.74	0.75	0.64	0.6	0.14	0.37	0.05	0.02
	Aux Pommes	0.83	0.79	0.89	0.78	0.04	0.05	0	0.53
Hydrotel1	Châteauguay	0.78	0.62	0.8	0.73	0.02	0.02	0	0
	Croche	0.86	0.79	0.88	0.76	0.15	0.07	0.4	0.05
	Chaudière	0.78	0.72	0.74	0.73	0.07	0.08	0.01	0.38
	Aux Brochets	0.69	0.66	0.71	0.7	0.03	0.06	0.05	0
	Boyer	0.76	0.69	0.73	0.68	0.03	0.09	0	0.81
	Aux Pommes*	0.78	0.68	0.78	0.71	0.04	0.04	0.91	0.01
Hydrotel2	Châteauguay	0.74	0.6	0.76	0.7	0.11	0.09	0.11	0
	Croche	0.84	0.77	0.83	0.74	0.16	0.18	0.56	0.1
	Chaudière	0.75	0.7	0.74	0.73	0.1	0.12	0.27	0.18
	Aux Brochets	0.65	0.62	0.65	0.66	0.04	0.07	0.34	0
	Boyer	0.74	0.69	0.72	0.67	0.04	0.07	0.02	0.69
	Aux Pommes*	0.78	0.68	0.78	0.71	0.04	0.04	0.91	0.01

Table-A I-3   Number of RHHUs for Hydotel 1 and Hydrotel 2 simulations as well as
number of rows and columns for WaSiM simulations

Catchment	Resolution (m)	RHHUs (Hydrotel 1)	RHHUs (Hydrotel 2)	Cloumns (WaSiM)	Rows (WaSiM)
	100	206	133	684	794
	250	204	97	273	318
Châteauguay	500	183	56	137	159
	1000	153	33	69	80
	100	117	87	389	862
	250	115	55	155	345
Croche	500	113	26	78	173
	1000	92	17	39	87
	100	97	41	393	438
	250	88	30	158	176
Chaudière	500	81	20	79	89
	1000	78	9	40	45
	100	55	27	311	366
	250	52	18	125	147
Aux Brochets	500	51	13	63	74
	1000	50	7	32	37
	100	17	7	175	222
	250	16	5	70	89
Boyer	500	15	5	35	45
	1000	13	3	18	23
	100	3	5	139	164
	250	1	3	56	66
Aux Pommes	500	1	1	29	33
	1000	1	1	14	17

Table-A I-4   Average annual precipitation (Pr) and mean temperature (Temp) for reference (ref, 2001-2020) and future (fut, 2081-2100) periods. The last column shows the number of ClimEx data-points inside the catchments

Catchment	Pr-fut (mm)	Pr-ref (mm)	Temp-fut (°C)	Temp (°C)	# Datapoints
Châteauguay	1135.1	1075.7	10.0	5.0	63
Croche	1129.5	1094.3	8.4	3.2	48
Chaudière	1195.9	1129.9	8.8	3.9	42
Aux Brochets	1174.3	1119.8	9.6	4.8	20
Boyer	1232.6	1199.1	10.2	5.0	6
Aux Pommes	1137.5	1101.4	8.5	3.3	4

BIBLIOGRAPHY

Aalbers, E. E., Lenderink, G., van Meijgaard, E. & van den Hurk, B. J. (2018). Local-scale changes in mean and heavy precipitation in Western Europe, climate change or internal variability? *Climate Dynamics*, 50(11), 4745–4766.

Aas, K., Czado, C., Frigessi, A. & Bakken, H. (2009). Pair-copula constructions of multiple dependence. *Insurance: Mathematics and economics*, 44(2), 182–198.

Abdulla, F. A. & Lettenmaier, D. P. (1997a). Application of regional parameter estimation schemes to simulate the water balance of a large continental river. *Journal of hydrology*, 197(1-4), 258–285.

Abdulla, F. A. & Lettenmaier, D. P. (1997b). Development of regional parameter estimation equations for a macroscale hydrologic model. *Journal of hydrology*, 197(1-4), 230–257.

Abrahart, R. J., Anctil, F., Coulibaly, P., Dawson, C. W., Mount, N. J., See, L. M., Shamseldin, A. Y., Solomatine, D. P., Toth, E. & Wilby, R. L. (2012). Two decades of anarchy? Emerging themes and outstanding challenges for neural network river forecasting. *Progress in Physical Geography*, 36(4), 480–513.

Abramowitz, G. (2005). Towards a benchmark for land surface models. *Geophysical Research Letters*, 32(22).

Addor, N., Rössler, O., Köplin, N., Huss, M., Weingartner, R. & Seibert, J. (2014). Robust changes and sources of uncertainty in the projected hydrological regimes of Swiss catchments. *Water Resources Research*, 50(10), 7541–7562. Retrieved from: https://doi.org/10.1002/2014WR015549.

Addor, N., Nearing, G., Prieto, C., Newman, A., Le Vine, N. & Clark, M. P. (2018). A ranking of hydrological signatures based on their predictability in space. *Water Resources Research*, 54(11), 8792–8812.

Ajami, N. K., Gupta, H., Wagener, T. & Sorooshian, S. (2004). Calibration of a semi-distributed hydrologic model for streamflow estimation along a river system. *Journal of hydrology*, 298(1-4), 112–135.

Alaoui, A., Rogger, M., Peth, S. & Blöschl, G. (2018). Does soil compaction increase floods? A review. *Journal of hydrology*, 557, 631–642.

Alexander, L. V., Zhang, X., Peterson, T. C., Caesar, J., Gleason, B., Klein Tank, A., Haylock, M., Collins, D., Trewin, B., Rahimzadeh, F. et al. (2006). Global observed changes in daily climate extremes of temperature and precipitation. *Journal of Geophysical Research: Atmospheres*, 111(D5).

Alfieri, L., Bisselink, B., Dottori, F., Naumann, G., de Roo, A., Salamon, P., Wyser, K. & Feyen, L. (2017). Global projections of river flood risk in a warmer world. *Earth's Future*, 5(2), 171–182.

Ali, A. S. A., Ebrahimi, S., Ashiq, M. M., Alasta, M. S. & Azari, B. (2022). CNN-Bi LSTM neural network for simulating groundwater level. *Environ Eng*, 8, 1–7.

Ali, R., Kuriqi, A., Abubaker, S. & Kisi, O. (2019). Long-term trends and seasonality detection of the observed flow in Yangtze River using Mann-Kendall and Sen's innovative trend method. *Water*, 11(9), 1855.

Anderson, G. J., Lucas, D. D. & Bonfils, C. (2018). Uncertainty Analysis of Simulations of the Turn-of-the-Century Drought in the Western United States. *Journal of Geophysical Research: Atmospheres*, 123(23), 13–219.

Arnold, J. G., Srinivasan, R., Muttiah, R. S. & Williams, J. R. (1998). Large area hydrologic modeling and assessment part I: model development 1. *JAWRA Journal of the American Water Resources Association*, 34(1), 73–89. Retrieved from: https://doi.org/10.1111/j.1752-1688.1998.tb05961.x.

Arsenault, R., Poulin, A., Côté, P., Brissette, F. et al. (2014). Comparison of stochastic optimization algorithms in hydrological model calibration. *J. Hydrol. Eng*, 19(7), 1374–1384.

Arsenault, R., Breton-Dufour, M., Poulin, A., Dallaire, G. & Romero-Lopez, R. (2019). Streamflow prediction in ungauged basins: analysis of regionalization methods in a hydrologically heterogeneous region of Mexico. *Hydrological Sciences Journal*, 64(11), 1297–1311.

Arsenault, R., Martel, J.-L., Brunet, F., Brissette, F. & Mai, J. (2022). Continuous streamflow prediction in ungauged basins: Long Short-Term Memory Neural Networks clearly outperform hydrological models. *Hydrology and Earth System Sciences Discussions*, 1–29.

Bachmair, S., Svensson, C., Hannaford, J., Barker, L. & Stahl, K. (2016). A quantitative analysis to objectively appraise drought indicators and model drought impacts. *Hydrology and Earth System Sciences*, 20(7), 2589–2609.

Bajamgnigni Gbambie, A. S., Poulin, A., Boucher, M.-A. & Arsenault, R. (2017). Added value of alternative information in interpolated precipitation datasets for hydrology. *Journal of Hydrometeorology*, 18(1), 247–264. Retrieved from: https://doi.org/10.5194/os-15-831-2019.

Balázs, B., Bíró, T., Dyke, G., Singh, S. K. & Szabó, S. (2018). Extracting water-related features using reflectance data and principal component analysis of Landsat images. *Hydrological Sciences Journal*, 63(2), 269–284.

Ban, N., Schmidli, J. & Schär, C. (2014). Cloud-resolving regional climate modeling approach in decade-long simulations. *21st Century Challenges in Regional Climate Modelling*, 1, 86.

Ban, N., Caillaud, C., Coppola, E., Pichelli, E., Sobolowski, S., Adinolfi, M., Ahrens, B., Alias, A., Anders, I., Bastin, S. et al. (2021). The first multi-model ensemble of regional climate simulations at kilometer-scale resolution, part I: evaluation of precipitation. *Climate Dynamics*, 57(1), 275–302.

Barendrecht, M. H., Viglione, A. & Blöschl, G. (2017). A dynamic framework for flood risk. *Water Security*, 1, 3–11.

Barnes, E. A. & Barnes, R. J. (2015). Estimating linear trends: simple linear regression versus epoch differences. *Journal of Climate*, 28(24), 9969–9976.

Baroni, G., Zink, M., Kumar, R., Samaniego, L. & Attinger, S. (2017). Effects of uncertainty in soil properties on simulated hydrological states and fluxes at different spatio-temporal scales. *Hydrology and Earth System Sciences*, 21(5), 2301–2320.

Barrios, M. & Francés, F. (2012). Spatial scale effect on the upper soil effective parameters of a distributed hydrological model. *Hydrological processes*, 26(7), 1022–1033. Retrieved from: https://doi.org/10.1002/hyp.8193.

Bates, B., Kundzewicz, Z. & Wu, S. (2008). *Climate change and water*. Intergovernmental Panel on Climate Change Secretariat.

Bathurst, J., Ewen, J., Parkin, G., O'Connell, P. & Cooper, J. (2004). Validation of catchment models for predicting land-use and climate change impacts. 3. Blind validation for internal and outlet responses. *Journal of Hydrology*, 287(1-4), 74–94.

Behrouz, M. S., Zhu, Z., Matott, L. S. & Rabideau, A. J. (2020). A new tool for automatic calibration of the Storm Water Management Model (SWMM). *Journal of Hydrology*, 581, 124436.

Bell, V., Kay, A., Jones, R., Moore, R. & Reynard, N. (2009). Use of soil data in a grid-based hydrological model to estimate spatial variation in changing flood risk across the UK. *Journal of Hydrology*, 377(3-4), 335–350.

Belmonte Rivas, M. & Stoffelen, A. (2019). Characterizing ERA-Interim and ERA5 surface wind biases using ASCAT. *Ocean Science*, 15(3). Retrieved from: https://doi.org/10.5194/os-15-831-2019.

Bender, J., Wahl, T., Müller, A. & Jensen, J. (2016). A multivariate design framework for river confluences. *Hydrological Sciences Journal*, 61(3), 471–482.

Bennett, B., Leonard, M., Deng, Y. & Westra, S. (2018). An empirical investigation into the effect of antecedent precipitation on flood volume. *Journal of hydrology*, 567, 435–445. doi: https://doi.org/10.1016/j.jhydrol.2018.10.025.

Berg, A. & Sheffield, J. (2018). Climate change and drought: the soil moisture perspective. *Current Climate Change Reports*, 4(2), 180–191.

Berg, P. & Haerter, J. (2013). Unexpected increase in precipitation intensity with temperature—A result of mixing of precipitation types? *Atmospheric Research*, 119, 56–61.

Bergström, S. et al. (1995). The HBV model. *Computer models of watershed hydrology.*, 443–476. Retrieved from: https://scholar.google.ca/scholar?hl=en&as_sdt=0%2C5&as_vis=1&q=The+HBV+model&btnG=.

Bertola, M., Viglione, A., Lun, D., Hall, J. & Blöschl, G. (2020). Flood trends in Europe: are changes in small and big floods different? *Hydrology and Earth System Sciences*, 24(4), 1805–1822.

Bertola, M., Viglione, A., Vorogushyn, S., Lun, D., Merz, B. & Blöschl, G. (2021). Do small and large floods have the same drivers of change? A regional attribution analysis in Europe. *Hydrology and Earth System Sciences*, 25(3), 1347–1364.

Best, M. J., Abramowitz, G., Johnson, H., Pitman, A., Balsamo, G., Boone, A., Cuntz, M., Decharme, B., Dirmeyer, P., Dong, J. et al. (2015). The plumbing of land surface models: benchmarking model performance. *Journal of Hydrometeorology*, 16(3), 1425–1442.

Bevacqua, E., Maraun, D., Hobæk Haff, I., Widmann, M. & Vrac, M. (2017). Multivariate statistical modelling of compound events via pair-copula constructions: analysis of floods in Ravenna (Italy). *Hydrology and Earth System Sciences*, 21(6), 2701–2723.

Beven, K. (1989). Changing ideas in hydrology—the case of physically-based models. *Journal of hydrology*, 105(1-2), 157–172.

Beven, K. (1995). Linking parameters across scales: subgrid parameterizations and scale dependent hydrological models. *Hydrological processes*, 9(5-6), 507–525. Retrieved from: https://doi.org/10.1002/hyp.3360090504.

Beven, K. (1997). TOPMODEL: a critique. *Hydrological processes*, 11(9), 1069–1085.

Beven, K. (2006). A manifesto for the equifinality thesis. *Journal of hydrology*, 320(1-2), 18–36. Retrieved from: https://doi.org/10.1016/j.jhydrol.2005.07.007.

Beven, K. (2010). *Environmental modelling: An uncertain future?* CRC Press.

Beven, K. (2016). Facets of uncertainty: epistemic uncertainty, non-stationarity, likelihood, hypothesis testing, and communication. *Hydrological Sciences Journal*, 61(9), 1652–1665.

Beven, K. (2019). How to make advances in hydrological modelling. *Hydrology Research*, 50(6), 1481–1494.

Beven, K. & Binley, A. (1992). The future of distributed models: model calibration and uncertainty prediction. *Hydrological processes*, 6(3), 279–298.

Beven, K. & Freer, J. (2001). Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *Journal of hydrology*, 249(1-4), 11–29.

Beven, K. J. (2000). Uniqueness of place and process representations in hydrological modelling. *Hydrology and earth system sciences*, 4(2), 203–213. Retrieved from: https://doi.org/10.5194/hess-4-203-2000.

Beven, K. J. (2011). *Rainfall-runoff modelling: the primer*. John Wiley & Sons. Retrieved from: https://www.wiley.com/en-ca/Rainfall+Runoff+Modelling%3A+The+Primer%2C+2nd+Edition-p-9781119951018.

Bhuiyan, A. et al. (2013). Investigation of Changes in Hydrological Processes using a Regional Climate Model.

Bierkens, M. F. (2015). Global hydrology 2015: State, trends, and directions. *Water Resources Research*, 51(7), 4923–4947.

Bishop, C. M. & Nasrabadi, N. M. (2006). *Pattern recognition and machine learning*. Springer.

Bjerknes, V. (1910). *Dynamic meteorology and hydrography*. Carnegie.

Blasone, R.-S., Vrugt, J. A., Madsen, H., Rosbjerg, D., Robinson, B. A. & Zyvoloski, G. A. (2008). Generalized likelihood uncertainty estimation (GLUE) using adaptive Markov Chain Monte Carlo sampling. *Advances in Water Resources*, 31(4), 630–648.

Blazejewski, R., Pilarczyk, K. & Przedwojski, B. (1995). *River training techniques: fundamentals, design and applications*. CRC Press.

Blöschl, G. (2022a). Flood generation: process patterns from the raindrop to the ocean. *Hydrology and Earth System Sciences*, 26(9), 2469–2480.

Blöschl, G. (2022b). Three hypotheses on changing river flood hazards. *Hydrology and Earth System Sciences Discussions*, 1–37.

Blöschl, G. & Sivapalan, M. (1995). Scale issues in hydrological modelling: a review. *Hydrological processes*, 9(3-4), 251–290.

Blöschl, G., Reszler, C. & Komma, J. (2008). A spatially distributed flash flood forecasting model. *Environmental Modelling & Software*, 23(4), 464–478. Retrieved from: https://doi.org/10.1016/j.envsoft.2007.06.010.

Blöschl, G., Bloschl, G., Sivapalan, M., Wagener, T., Savenije, H. & Viglione, A. (2013). *Runoff prediction in ungauged basins: synthesis across processes, places and scales*. Cambridge University Press.

Blöschl, G., Blaschke, A., Broer, M., Bucher, C., Carr, G., Chen, X., Eder, A., Exner-Kittridge, M., Farnleitner, A., Flores-Orozco, A. et al. (2016). The hydrological open air laboratory (HOAL) in Petzenkirchen: A hypothesis-driven observatory. *Hydrology and Earth System Sciences*, 20(1), 227–255.

Blöschl, G., Hall, J., Parajka, J., Perdigão, R. A., Merz, B., Arheimer, B., Aronica, G. T., Bilibashi, A., Bonacci, O., Borga, M. et al. (2017). Changing climate shifts timing of European floods. *Science*, 357(6351), 588–590.

Blöschl, G., Bierkens, M. F., Chambel, A., Cudennec, C., Destouni, G., Fiori, A., Kirchner, J. W., McDonnell, J. J., Savenije, H. H., Sivapalan, M. et al. (2019). Twenty-three unsolved problems in hydrology (UPH)–a community perspective. *Hydrological Sciences Journal*, 64(10), 1141–1158. Retrieved from: https://doi.org/10.1080/02626667.2019.1620507.

Bormann, H., Breuer, L., Gräff, T., Huisman, J. & Croke, B. (2009). Assessing the impact of land use change on hydrology by ensemble modelling (LUCHEM) IV: Model sensitivity to data aggregation and spatial (re-) distribution. *Advances in water resources*, 32(2), 171–192. Retrieved from: https://doi.org/10.1016/j.advwatres.2008.01.002.

Bosshard, T., Carambia, M., Goergen, K., Kotlarski, S., Krahe, P., Zappa, M. & Schär, C. (2013). Quantifying uncertainty sources in an ensemble of hydrological climate-impact projections. *Water Resources Research*, 49(3), 1523–1536.

Boucher, M.-A., Quilty, J. & Adamowski, J. (2020). Data assimilation for streamflow forecasting using extreme learning machines and multilayer perceptrons. *Water Resources Research*, 56(6), e2019WR026226.

Boulesteix, A.-L., Janitza, S., Kruppa, J. & König, I. R. (2012). Overview of random forest methodology and practical guidance with emphasis on computational biology and bioinformatics. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 2(6), 493–507.

Bowles, D. S. & O'Connell, P. E. (2012). *Recent advances in the modeling of hydrologic systems*. Springer Science & Business Media.

Brakenridge, G., Syvitski, J., Overeem, I., Higgins, S., Kettner, A., Stewart-Moore, J. & Westerhoff, R. (2013). Global mapping of storm surges and the assessment of coastal vulnerability. *Natural hazards*, 66(3), 1295–1312.

Branstator, G. & Teng, H. (2010). Two limits of initial-value decadal predictability in a CGCM. *Journal of climate*, 23(23), 6292–6311.

Branstator, G., Teng, H., Meehl, G. A., Kimoto, M., Knight, J. R., Latif, M. & Rosati, A. (2012). Systematic estimates of initial-value decadal predictability for six AOGCMs. *Journal of Climate*, 25(6), 1827–1846.

Breiman, L. (2001). Random forests. *Machine learning*, 45(1), 5–32.

Breinl, K., Lun, D., Müller-Thomy, H. & Blöschl, G. (2021). Understanding the relationship between rainfall and flood probabilities through combined intensity-duration-frequency analysis. *Journal of Hydrology*, 602, 126759.

Bresson, É., Laprise, R., Paquin, D., Thériault, J. & de Elía, R. (2017). Evaluating the ability of CRCM5 to simulate mixed precipitation. *Atmosphere-Ocean*, 55(2), 79–93.

Breuer, L., Huisman, J., Willems, P., Bormann, H., Bronstert, A., Croke, B., Frede, H.-G., Gräff, T., Hubrechts, L., Jakeman, A. et al. (2009). Assessing the impact of land use change on hydrology by ensemble modeling (LUCHEM). I: Model intercomparison with current land use. *Advances in Water Resources*, 32(2), 129–146. Retrieved from: https://doi.org/10.1016/j.advwatres.2008.10.003.

Bronstert, A., Bárdossy, A., Bismuth, C., Buiteveld, H., Disse, M., Engel, H., Fritsch, U., Hundecha, Y., Lammersen, R., Niehoff, D. et al. (2007). Multi-scale modelling of land-use change and river training effects on floods in the Rhine basin. *River Research and Applications*, 23(10), 1102–1125.

Brown, A. E., Zhang, L., McMahon, T. A., Western, A. W. & Vertessy, R. A. (2005). A review of paired catchment studies for determining changes in water yield resulting from alterations in vegetation. *Journal of hydrology*, 310(1-4), 28–61.

Brutsaert, W. et al. (2005). *Hydrology: an introduction*. Cambridge university press.

Brynjarsdottir, J. & OHagan, A. (2014). Learning about physical parameters: The importance of model discrepancy. *Inverse problems*, 30(11), 114007. Retrieved from: http://dx.doi.org/10.1088/0266-5611/30/11/114007.

Buizza, R., Milleer, M. & Palmer, T. N. (1999). Stochastic representation of model uncertainties in the ECMWF ensemble prediction system. *Quarterly Journal of the Royal Meteorological Society*, 125(560), 2887–2908. Retrieved from: https://doi.org/10.1002/qj.49712556006.

Buonomo, E., Jones, R., Huntingford, C. & Hannaford, J. (2007). On the robustness of changes in extreme precipitation over Europe from two high resolution climate change simulations. *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography*, 133(622), 65–81.

Burn, D. H. & Elnur, M. A. H. (2002). Detection of hydrologic trends and variability. *Journal of hydrology*, 255(1-4), 107–122.

Buttinger-Kreuzhuber, A., Horváth, Z., Noelle, S., Blöschl, G. & Waser, J. (2019). A fast second-order shallow water scheme on two-dimensional structured grids over abrupt topography. *Advances in water resources*, 127, 89–108.

Buttle, J. M., Allen, D. M., Caissie, D., Davison, B., Hayashi, M., Peters, D. L., Pomeroy, J. W., Simonovic, S., St-Hilaire, A. & Whitfield, P. H. (2016). Flood processes in Canada: Regional and special aspects. *Canadian Water Resources Journal/Revue canadienne des ressources hydriques*, 41(1-2), 7–30.

Butts, M. B., Payne, J. T., Kristensen, M. & Madsen, H. (2004). An evaluation of the impact of model structure on hydrological modelling uncertainty for streamflow simulation. *Journal of hydrology*, 298(1-4), 242–266. Retrieved from: https://doi.org/10.1016/j.jhydrol.2004.03.042.

Buytaert, W. & Beven, K. (2011). Models as multiple working hypotheses: hydrological simulation of tropical alpine wetlands. *Hydrological Processes*, 25(11), 1784–1799.

Cannon, A. J. (2018). Multivariate quantile mapping bias correction: an N-dimensional probability density function transform for climate model simulations of multiple variables. *Climate dynamics*, 50(1), 31–49.

Cannon, A. J., Piani, C. & Sippel, S. (2020). Bias correction of climate model output for impact models. In *Climate Extremes and Their Implications for Impact and Risk Assessment* (pp. 77–104). Elsevier.

Cao, X., Lyu, H., Ni, G., Tian, F., Ma, Y. & Grimmond, C. (2020a). Spatial Scale Effect of Surface Routing and Its Parameter Upscaling for Urban Flood Simulation Using a Grid-Based Model. *Water Resources Research*, 56(2), e2019WR025468. Retrieved from: https://doi.org/10.1029/2019WR025468.

Cao, X., Ni, G., Qi, Y. & Liu, B. (2020b). Does Subgrid Routing Information Matter for Urban Flood Forecasting? A Multiscenario Analysis at the Land Parcel Scale. *Journal of Hydrometeorology*, 21(9), 2083–2099. Retrieved from: https://doi.org/10.1175/JHM-D-20-0075.1.

Chen, J., Brissette, F. P., Poulin, A. & Leconte, R. (2011). Overall uncertainty study of the hydrological impacts of climate change for a Canadian watershed. *Water Resources Research*, 47(12).

Chen, J., Brissette, F. P., Chaumont, D. & Braun, M. (2013). Finding appropriate bias correction methods in downscaling precipitation for hydrologic impact studies over North America. *Water Resources Research*, 49(7), 4187–4205.

Chen, J., Brissette, F. P., Lucas-Picher, P. & Caya, D. (2017). Impacts of weighting climate models for hydro-meteorological climate change studies. *Journal of hydrology*, 549, 534–546.

Chen, J., Chen, H. & Guo, S. (2018). Multi-site precipitation downscaling using a stochastic weather generator. *Climate dynamics*, 50(5), 1975–1992.

Chen, J., Brissette, F. P., Zhang, X. J., Chen, H., Guo, S. & Zhao, Y. (2019a). Bias correcting climate model multi-member ensembles to assess climate change impacts on hydrology. *Climatic Change*, 153(3), 361–377.

Chen, Y., Xu, C.-Y., Chen, X., Xu, Y., Yin, Y., Gao, L. & Liu, M. (2019b). Uncertainty in simulation of land-use change impacts on catchment runoff with multi-timescales based on the comparison of the HSPF and SWAT models. *Journal of Hydrology*, 573, 486–500. Retrieved from: https://doi.org/10.1016/j.jhydrol.2019.03.091.

Chlumsky, R., Mai, J., Craig, J. R. & Tolson, B. A. (2021). Simultaneous calibration of hydrologic model structure and parameters using a blended model. *Water Resources Research*, 57(5), e2020WR029229.

Chow, V. T., Maidment, D. R. & Mays, L. W. (1988). *Applied hydrology*.

Clark, M. P., Kavetski, D. & Fenicia, F. (2011). Pursuing the method of multiple working hypotheses for hydrological modeling. *Water Resources Research*, 47(9).

Clark, M. P., Nijssen, B., Lundquist, J. D., Kavetski, D., Rupp, D. E., Woods, R. A., Freer, J. E., Gutmann, E. D., Wood, A. W., Brekke, L. D. et al. (2015). A unified approach for process-based hydrologic modeling: 1. Modeling concept. *Water Resources Research*, 51(4), 2498–2514. Retrieved from: https://doi.org/10.1002/2015WR017200.

Clark, M. P., Bierkens, M. F., Samaniego, L., Woods, R. A., Uijlenhoet, R., Bennett, K. E., Pauwels, V., Cai, X., Wood, A. W. & Peters-Lidard, C. D. (2017). The evolution of process-based hydrologic models: historical challenges and the collective quest for physical realism. *Hydrology and Earth System Sciences (Online)*, 21(LA-UR-17-27603). Retrieved from: https://doi.org/10.5194/hess-21-3427-2017.

Cloke, H. & Pappenberger, F. (2009). Ensemble flood forecasting: A review. *Journal of hydrology*, 375(3-4), 613–626. Retrieved from: https://doi.org/10.1016/j.jhydrol.2009.06.005.

Cole, S., Moore, R., Bell, V. & Jones, D. (2006). Issues in flood forecasting: ungauged basins, extreme floods and uncertainty. *Frontiers in Flood Forecasting, 8th Kovacs Colloquium, UNESCO, Paris*, 305, 103–122.

Coles, S., Bawa, J., Trenner, L. & Dorazio, P. (2001). *An introduction to statistical modeling of extreme values*. Springer. Retrieved from: https://link.springer.com/book/10.1007%2F978-1-4471-3675-0.

Coles, S. G. & Tawn, J. A. (1991). Modelling extreme multivariate events. *Journal of the Royal Statistical Society: Series B (Methodological)*, 53(2), 377–392.

Contractor, S., Donat, M. G. & Alexander, L. V. (2021). Changes in observed daily precipitation over global land areas since 1950. *Journal of Climate*, 34(1), 3–19.

Cook, R. D. & Weisberg, S. (1982). *Residuals and influence in regression*. New York: Chapman and Hall.

Cordery, I. (1970). Antecedent wetness for design flood estimation. *Civil Engineering Transaction*.

Cossart, É. & Fressard, M. (2017). Assessment of structural sediment connectivity within catchments: insights from graph theory. *Earth Surface Dynamics*, 5(2), 253–268.

Craig, J. R., Brown, G., Chlumsky, R., Jenkinson, W., Jost, G., Lee, K., Mai, J., Serrer, M., Snowdon, A. P., Sgro, N. et al. (2020). Flexible watershed simulation with the Raven hydrological modelling framework. *Environmental Modelling & Software*, 104728. Retrieved from: https://doi.org/10.1016/j.envsoft.2020.104728.

Cressie, N. (1990). The origins of kriging. *Mathematical geology*, 22(3), 239–252.

Croke, J., Mockler, S., Fogarty, P. & Takken, I. (2005). Sediment concentration changes in runoff pathways from a forest road network and the resultant spatial pattern of catchment connectivity. *Geomorphology*, 68(3-4), 257–268.

Danielson, J. J. & Gesch, D. B. (2011). *Global multi-resolution terrain elevation data 2010 (GMTED2010)*. US Department of the Interior, US Geological Survey Washington, DC, USA.

Dankers, R., Arnell, N. W., Clark, D. B., Falloon, P. D., Fekete, B. M., Gosling, S. N., Heinke, J., Kim, H., Masaki, Y., Satoh, Y. et al. (2014). First look at changes in flood hazard in the Inter-Sectoral Impact Model Intercomparison Project ensemble. *Proceedings of the National Academy of Sciences*, 111(9), 3257–3261.

Darbandsari, P. & Coulibaly, P. (2020). Inter-comparison of lumped hydrological models in data-scarce watersheds using different precipitation forcing data sets: Case study of Northern Ontario, Canada. *Journal of Hydrology: Regional Studies*, 31, 100730.

Das, T., Bárdossy, A., Zehe, E. & He, Y. (2008). Comparison of conceptual model performance using different representations of spatial variability. *Journal of Hydrology*, 356(1-2), 106–118.

Davison, A. C., Padoan, S. A., Ribatet, M. et al. (2012). Statistical modeling of spatial extremes. *Statistical science*, 27(2), 161–186.

Dawson, C. & Wilby, R. (2001). Hydrological modelling using artificial neural networks. *Progress in physical Geography*, 25(1), 80–108.

Dembélé, M., Hrachowitz, M., Savenije, H. H., Mariéthoz, G. & Schaefli, B. (2020). Improving the Predictive Skill of a Distributed Hydrological Model by Calibration on Spatial Patterns With Multiple Satellite Data Sets. *Water resources research*, 56(1), e2019WR026085. Retrieved from: https://doi.org/10.1029/2019WR026085.

Deser, C., Knutti, R., Solomon, S. & Phillips, A. S. (2012a). Communication of the role of natural variability in future North American climate. *Nature Climate Change*, 2(11), 775.

Deser, C., Phillips, A., Bourdette, V. & Teng, H. (2012b). Uncertainty in climate change projections: the role of internal variability. *Climate dynamics*, 38(3-4), 527–546.

Detty, J. & McGuire, K. (2010). Topographic controls on shallow groundwater dynamics: implications of hydrologic connectivity between hillslopes and riparian zones in a till mantled catchment. *Hydrological Processes*, 24(16), 2222–2236.

Devia, G. K., Ganasri, B. & Dwarakish, G. (2015). A review on hydrological models. *Aquatic Procedia*, 4, 1001–1007. Retrieved from: https://doi.org/10.1016/j.aqpro.2015.02.126.

Di Baldassarre, G., Viglione, A., Carr, G., Kuil, L., Yan, K., Brandimarte, L. & Blöschl, G. (2015). Debates—Perspectives on socio-hydrology: Capturing feedbacks between physical and social processes. *Water Resources Research*, 51(6), 4770–4781.

Dickinson, R. E., Errico, R. M., Giorgi, F. & Bates, G. T. (1989). A regional climate model for the western United States. *Climatic change*, 15(3), 383–422.

Ding, J., Wallner, M., Müller, H. & Haberlandt, U. (2016). Estimation of instantaneous peak flows from maximum mean daily flows using the HBV hydrological model. *Hydrological Processes*, 30(9), 1431–1448.

Ding, Q. & Steig, E. J. (2013). Temperature change on the Antarctic Peninsula linked to the tropical Pacific. *Journal of Climate*, 26(19), 7570–7585.

Dingman, S. (2002). Physical Hydrology Waveland Press. *Long Grove, Illinois*.

Dixon, B. & Earls, J. (2012). Effects of urbanization on streamflow using SWAT with real and simulated meteorological data. *Applied geography*, 35(1-2), 174–190. Retrieved from: https://doi.org/10.1016/j.apgeog.2012.06.010.

Do, H. X., Zhao, F., Westra, S., Leonard, M., Gudmundsson, L., Boulange, J. E. S., Chang, J., Ciais, P., Gerten, D., Gosling, S. N. et al. (2020). Historical and future changes in global flood magnitude–evidence from a model–observation investigation. *Hydrology and Earth System Sciences*, 24(3), 1543–1564.

Donat, M. G., Lowry, A. L., Alexander, L. V., O'Gorman, P. A. & Maher, N. (2016). More extreme precipitation in the world's dry and wet regions. *Nature Climate Change*, 6(5), 508.

Donat, M., Alexander, L. V., Yang, H., Durre, I., Vose, R., Dunn, R. J., Willett, K. M., Aguilar, E., Brunet, M., Caesar, J. et al. (2013). Updated analyses of temperature and precipitation extreme indices since the beginning of the twentieth century: The HadEX2 dataset. *Journal of Geophysical Research: Atmospheres*, 118(5), 2098–2118.

Dooge, J. C. (1986). Looking for hydrologic laws. *Water Resources Research*, 22(9S), 46S–58S. Retrieved from: https://doi.org/10.1029/WR022i09Sp0046S.

Dottori, F., Szewczyk, W., Ciscar, J.-C., Zhao, F., Alfieri, L., Hirabayashi, Y., Bianchi, A., Mongelli, I., Frieler, K., Betts, R. A. et al. (2018). Increased human and economic losses from river flooding with anthropogenic warming. *Nature Climate Change*, 8(9), 781–786.

Drobinski, P., Da Silva, N., Bastin, S., Mailler, S., Muller, C., Ahrens, B., Christensen, O. B. & Lionello, P. (2020). How warmer and drier will the Mediterranean region be at the end of the twenty-first century? *Regional Environmental Change*, 20(3), 1–12.

Duan, Q., Schaake, J., Andréassian, V., Franks, S., Goteti, G., Gupta, H., Gusev, Y., Habets, F., Hall, A., Hay, L. et al. (2006). Model Parameter Estimation Experiment (MOPEX): An overview of science strategy and major results from the second and third workshops. *Journal of Hydrology*, 320(1-2), 3–17.

Duan, Q., Gupta, V. K. & Sorooshian, S. (1993). Shuffled complex evolution approach for effective and efficient global minimization. *Journal of optimization theory and applications*, 76(3), 501–521.

Duan, S. & Ullrich, P. (2021). A comprehensive investigation of machine learning models for estimating daily snow water equivalent over the Western US.

Duethmann, D., Blöschl, G. & Parajka, J. (2020). Why does a conceptual hydrological model fail to correctly predict discharge changes in response to climate change? *Hydrology and Earth System Sciences*, 24(7), 3493–3511.

Dunne, T. & Black, R. D. (1970). Partial area contributions to storm runoff in a small New England watershed. *Water resources research*, 6(5), 1296–1311.

Edwards, P. N. (1999). Global climate science, uncertainty and politics: Data-laden models, model-filtered data. *Science as culture*, 8(4), 437–472.

Edwards, P. N. (2011). History of climate modeling. *Wiley Interdisciplinary Reviews: Climate Change*, 2(1), 128–139.

Eiriksson, D., Whitson, M., Luce, C. H., Marshall, H. P., Bradford, J., Benner, S. G., Black, T., Hetrick, H. & McNamara, J. P. (2013). An evaluation of the hydrologic relevance of lateral flow in snow at hillslope and catchment scales. *Hydrological Processes*, 27(5), 640–654.

Euser, T., Winsemius, H., Hrachowitz, M., Fenicia, F., Uhlenbrook, S. & Savenije, H. (2013). A framework to assess the realism of model structures using hydrological signatures. *Hydrology and Earth System Sciences, 17 (5), 2013*. Retrieved from: https://doi.org/10.5194/hess-17-1893-2013.

Euser, T., Hrachowitz, M., Winsemius, H. C. & Savenije, H. H. (2015). The effect of forcing and landscape distribution on performance and consistency of model structures. *Hydrological Processes*, 29(17), 3727–3743.

Faghih, M., Brissette, F. & Sabeti, P. (2022). Impact of correcting sub-daily climate model biases for hydrological studies. *Hydrology and Earth System Sciences*, 26(6), 1545–1563.

Famiglietti, J. & Wood, E. F. (1995). Effects of spatial variability and scale on areally averaged evapotranspiration. *Water Resources Research*, 31(3), 699–712. Retrieved from: https://doi.org/10.1029/94WR02820.

Faramarzi, M., Abbaspour, K. C., Vaghefi, S. A., Farzaneh, M. R., Zehnder, A. J., Srinivasan, R. & Yang, H. (2013). Modeling impacts of climate change on freshwater availability in Africa. *Journal of Hydrology*, 480, 85–101. Retrieved from: https://doi.org/10.1016/j.jhydrol.2012.12.016.

Fatichi, S., Vivoni, E. R., Ogden, F. L., Ivanov, V. Y., Mirus, B., Gochis, D., Downer, C. W., Camporese, M., Davison, J. H., Ebel, B. et al. (2016). An overview of current applications, challenges, and future trends in distributed process-based models in hydrology. *Journal of Hydrology*, 537, 45–60. Retrieved from: https://doi.org/10.1016/j.jhydrol.2016.03.026.

Fenicia, F., Kavetski, D. & Savenije, H. H. (2011). Elements of a flexible approach for conceptual hydrological modeling: 1. Motivation and theoretical development. *Water Resources Research*, 47(11).

Field, C. B., Barros, V., Stocker, T. F. & Dahe, Q. (2012). *Managing the risks of extreme events and disasters to advance climate change adaptation: special report of the intergovernmental panel on climate change*. Cambridge University Press.

Fischer, E. M. & Knutti, R. (2014). Detection of spatially aggregated changes in temperature and precipitation extremes. *Geophysical Research Letters*, 41(2), 547–554.

Fischer, E. M., Beyerle, U. & Knutti, R. (2013). Robust spatially aggregated projections of climate extremes. *Nature Climate Change*, 3(12), 1033.

Fischer, E. M., Sedláček, J., Hawkins, E. & Knutti, R. (2014). Models agree on forced response pattern of precipitation and temperature extremes. *Geophysical Research Letters*, 41(23), 8554–8562.

Floods, J. & Floods, J.-S. C. (2004). Munich Re NatCatSERVICE®. *Munich Re: Munich, Germany*.

Flury, M., Flühler, H., Jury, W. A. & Leuenberger, J. (1994). Susceptibility of soils to preferential flow of water: A field study. *Water resources research*, 30(7), 1945–1954.

Foglia, L., Hill, M. C., Mehl, S. W. & Burlando, P. (2009). Sensitivity analysis, calibration, and testing of a distributed hydrological model using error-based weighting and one objective function. *Water Resources Research*, 45(6).

Foley, A. (2010). Uncertainty in regional climate modelling: A review. *Progress in Physical Geography*, 34(5), 647–670.

for Disaster Reduction. Secretariat, U. N. I. S. (2015). *Global Assessment Report on Disaster Risk Reduction 2015: Making Development Sustainable: the Future of Disaster Risk Management*. UN.

Fortin, J.-P., Turcotte, R., Massicotte, S., Moussa, R., Fitzback, J. & Villeneuve, J.-P. (2001a). Distributed watershed model compatible with remote sensing and GIS data. II: Application to Chaudière watershed. *Journal of Hydrologic Engineering*, 6(2), 100–108. Retrieved from: https://doi.org/10.1061/(ASCE)1084-0699(2001)6:2(91).

Fortin, J.-P., Turcotte, R., Massicotte, S., Moussa, R., Fitzback, J. & Villeneuve, J.-P. (2001b). Distributed watershed model compatible with remote sensing and GIS data. II: Application to Chaudière watershed. *Journal of Hydrologic Engineering*, 6(2), 100–108. Retrieved from: https://doi.org/10.1061/(ASCE)1084-0699(2001)6:2(100).

Fowler, H. J., Ali, H., Allan, R. P., Ban, N., Barbero, R., Berg, P., Blenkinsop, S., Cabi, N. S., Chan, S., Dale, M. et al. (2021). Towards advancing scientific knowledge of climate change impacts on short-duration rainfall extremes. *Philosophical Transactions of the Royal Society A*, 379(2195), 20190542.

Fowler, H., Ekström, M., Blenkinsop, S. & Smith, A. (2007). Estimating change in extreme European precipitation using a multimodel ensemble. *Journal of Geophysical Research: Atmospheres*, 112(D18).

Fox, J. & Weisberg, S. (2018). *An R companion to applied regression*. Sage Publications.

Francis, J. A. & Vavrus, S. J. (2012). Evidence linking Arctic amplification to extreme weather in mid-latitudes. *Geophysical research letters*, 39(6).

François, B., Schlef, K., Wi, S. & Brown, C. (2019). Design considerations for riverine floods in a changing climate–a review. *Journal of Hydrology*, 574, 557–573.

Freeze, R. A. & Harlan, R. (1969). Blueprint for a physically-based, digitally-simulated hydrologic response model. *Journal of hydrology*, 9(3), 237–258.

Gao, M., Li, H.-Y., Liu, D., Tang, J., Chen, X., Chen, X., Blöschl, G. & Leung, L. R. (2018). Identifying the dominant controls on macropore flow velocity in soils: A meta-analysis. *Journal of Hydrology*, 567, 590–604.

Garavaglia, F., Lay, M. L., Gottardi, F., Garçon, R., Gailhard, J., Paquet, E. & Mathevet, T. (2017). Impact of model structure on flow simulation and hydrological realism: from a lumped to a semi-distributed approach. *Hydrology and Earth System Sciences*, 21(8), 3937–3952.

Gauch, M., Kratzert, F., Klotz, D., Nearing, G., Lin, J. & Hochreiter, S. (2021a). Rainfall–runoff prediction at multiple timescales with a single Long Short-Term Memory network. *Hydrology and Earth System Sciences*, 25(4), 2045–2062.

Gauch, M., Mai, J. & Lin, J. (2021b). The proper care and feeding of CAMELS: How limited training data affects streamflow prediction. *Environmental Modelling & Software*, 135, 104926.

Gaur, A., Gaur, A. & Simonovic, S. (2018). Future Changes in Flood Hazards across Canada under a Changing Climate. *Water*, 10(10), 1441.

Gebremicael, T., Mohamed, Y. & Van der Zaag, P. (2019). Attributing the hydrological impact of different land use types and their long-term dynamics through combining parsimonious hydrological modelling, alteration analysis and PLSR analysis. *Science of the Total Environment*, 660, 1155–1167.

Giuntoli, I., Vidal, J.-P., Prudhomme, C. & Hannah, D. M. (2015). Future hydrological extremes: the uncertainty from multiple global climate and global hydrological models. *Earth System Dynamics*, 6(1), 267–285.

Gneiting, T. & Katzfuss, M. (2014). Probabilistic forecasting. *Annual Review of Statistics and Its Application*, 1, 125–151. Retrieved from: https://doi.org/10.1146/annurev-statistics-062713-085831.

Golden, H. E. & Knightes, C. D. (2011). Simulated watershed mercury and nitrate flux responses to multiple land cover conversion scenarios. *Environmental Toxicology and Chemistry*, 30(4), 773–786.

Gomi, T., Sidle, R. C. & Richardson, J. S. (2002). Understanding processes and downstream linkages of headwater systems: headwaters differ from downstream reaches by their close coupling to hillslope processes, more temporal and spatial variation, and their need for different means of protection from land use. *BioScience*, 52(10), 905–916.

Götzinger, J. & Bárdossy, A. (2007). Comparison of four regionalisation methods for a distributed hydrological model. *Journal of Hydrology*, 333(2-4), 374–384.

Grams, C. M., Binder, H., Pfahl, S., Piaget, N. & Wernli, H. (2014). Atmospheric processes triggering the central European floods in June 2013. *Natural Hazards and Earth System Sciences*, 14(7), 1691–1702.

Grey, D. & Sadoff, C. W. (2007). Sink or swim? Water security for growth and development. *Water policy*, 9(6), 545–571.

Griffis, V. & Stedinger, J. (2007). Log-Pearson type 3 distribution and its application in flood frequency analysis. I: Distribution characteristics. *Journal of Hydrologic Engineering*, 12(5), 482–491.

Grimaldi, S., Petroselli, A., Tauro, F. & Porfiri, M. (2012). Time of concentration: a paradox in modern hydrology. *Hydrological Sciences Journal*, 57(2), 217–228.

Guerreiro, S. B., Fowler, H. J., Barbero, R., Westra, S., Lenderink, G., Blenkinsop, S., Lewis, E. & Li, X.-F. (2018). Detection of continental-scale intensification of hourly rainfall extremes. *Nature Climate Change*, 8(9), 803–807.

Guo, Y., Zhang, Y., Zhang, L. & Wang, Z. (2021). Regionalization of hydrological modeling for predicting streamflow in ungauged catchments: A comprehensive review. *Wiley Interdisciplinary Reviews: Water*, 8(1), e1487.

Gupta, A. & Govindaraju, R. (2019). Propagation of structural uncertainty in watershed hydrologic models. *Journal of Hydrology*, 575, 66–81.

Gupta, H. V., Kling, H., Yilmaz, K. K. & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of hydrology*, 377(1-2), 80–91. Retrieved from: https://doi.org/10.1016/j.jhydrol.2009.08.003.

Gupta, V. K., Rodríguez-Iturbe, I. & Wood, E. F. (2012). *Scale problems in hydrology: runoff generation and basin response*. Springer Science & Business Media.

Gurney, K. R., Kılkış, Ş., Seto, K. C., Lwasa, S., Moran, D., Riahi, K., Keller, M., Rayner, P. & Luqman, M. (2022). Greenhouse gas emissions from global cities under SSP/RCP scenarios, 1990 to 2100. *Global Environmental Change*, 73, 102478.

Gutowski, W. J., Ullrich, P. A., Hall, A., Leung, L. R., O'Brien, T. A., Patricola, C. M., Arritt, R., Bukovsky, M., Calvin, K. V., Feng, Z. et al. (2020). The ongoing need for high-resolution regional climate models: Process understanding and stakeholder information. *Bulletin of the American Meteorological Society*, 101(5), E664–E683.

Haerter, J. O. & Berg, P. (2009). Unexpected rise in extreme precipitation caused by a shift in rain type? *Nature Geoscience*, 2(6), 372.

Haerter, J., Berg, P. & Hagemann, S. (2010). Heavy rain intensity distributions on varying time scales and at different temperatures. *Journal of Geophysical Research: Atmospheres*, 115(D17).

Hall, J., Arheimer, B., Borga, M., Brázdil, R., Claps, P., Kiss, A., Kjeldsen, T., Kriaučiūnienė, J., Kundzewicz, Z. W., Lang, M. et al. (2014). Understanding flood regime changes in Europe: a state-of-the-art assessment. *Hydrology and Earth System Sciences*, 18(7), 2735–2772.

Hallema, D. W., Moussa, R., Sun, G. & McNulty, S. G. (2016). Surface storm flow prediction on hillslopes based on topography and hydrologic connectivity. *Ecological Processes*, 5(1), 1–13.

Halverson, M. & Fleming, S. (2015). Complex network theory, streamflow, and hydrometric monitoring system design. *Hydrology and Earth System Sciences*, 19(7), 3301–3318.

Hamon, W. R. (1961). Estimating potential evapotranspiration. *Journal of the Hydraulics Division*, 87(3), 107–120. Retrieved from: https://doi.org/10.1061/JYCEAJ.0000599.

Hattermann, F., Vetter, T., Breuer, L., Su, B., Daggupati, P., Donnelly, C., Fekete, B., Flörke, F., Gosling, S. N., Hoffmann, P. et al. (2018). Sources of uncertainty in hydrological climate impact assessment: a cross-scale study. *Environmental Research Letters*, 13(1), 015006.

Heffernan, J. E. & Tawn, J. A. (2004). A conditional approach for multivariate extreme values (with discussion). *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 66(3), 497–546.

Hersbach, H. & Dee, D. (2016). ERA5 reanalysis is in production. *ECMWF newsletter*, 147(7), 5–6.

Heuvelmans, G., Muys, B. & Feyen, J. (2006). Regionalisation of the parameters of a hydrological model: Comparison of linear regression models with artificial neural nets. *Journal of Hydrology*, 319(1-4), 245–265.

Hill, C., DeLuca, C., Suarez, M., Da Silva, A. et al. (2004). The architecture of the earth system modeling framework. *Computing in Science & Engineering*, 6(1), 18.

Hill, M. C., Kavetski, D., Clark, M., Ye, M., Arabi, M., Lu, D., Foglia, L. & Mehl, S. (2016). Practical use of computationally frugal model analysis methods. *Groundwater*, 54(2), 159–170.

Hochreiter, S. & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8), 1735–1780.

Hodgkins, G. A., Whitfield, P. H., Burn, D. H., Hannaford, J., Renard, B., Stahl, K., Fleig, A. K., Madsen, H., Mediero, L., Korhonen, J. et al. (2017). Climate-driven variability in the occurrence of major floods across North America and Europe. *Journal of Hydrology*, 552, 704–717.

Hong, C.-C., Hsu, H.-H., Lin, N.-H. & Chiu, H. (2011). Roles of European blocking and tropical-extratropical interaction in the 2010 Pakistan flooding. *Geophysical Research Letters*, 38(13).

Horton, R. E. (1933). The role of infiltration in the hydrologic cycle. *Eos, Transactions American Geophysical Union*, 14(1), 446–460.

Horváth, Z., Buttinger-Kreuzhuber, A., Konev, A., Cornel, D., Komma, J., Blöschl, G., Noelle, S. & Waser, J. (2020). Comparison of fast shallow-water schemes on real-world floods. *Journal of Hydraulic Engineering*, 146(1), 05019005.

Hrachowitz, M. & Clark, M. P. (2017). HESS Opinions: The complementary merits of competing modelling philosophies in hydrology. *Hydrol. Earth Syst. Sci*, 21(8), 3953–3973. Retrieved from: https://doi.org/10.5194/hess-21-3953-2017.

Hrachowitz, M., Savenije, H., Blöschl, G., McDonnell, J., Sivapalan, M., Pomeroy, J., Arheimer, B., Blume, T., Clark, M., Ehret, U. et al. (2013). A decade of Predictions in Ungauged Basins (PUB)—a review. *Hydrological sciences journal*, 58(6), 1198–1255. Retrieved from: https://doi.org/10.1080/02626667.2013.803183.

Hsu, K.-l., Gupta, H. V. & Sorooshian, S. (1995). Artificial neural network modeling of the rainfall-runoff process. *Water resources research*, 31(10), 2517–2530.

Huisman, J., Breuer, L., Bormann, H., Bronstert, A., Croke, B., Frede, H.-G., Gräff, T., Hubrechts, L., Jakeman, A., Kite, G. et al. (2009). Assessing the impact of land use change on hydrology by ensemble modeling (LUCHEM) III: Scenario analysis. *Advances in Water Resources*, 32(2), 159–170. Retrieved from: https://doi.org/10.1016/j.advwatres.2008. 06.009.

Hundecha, Y. & Bárdossy, A. (2004). Modeling of the effect of land use changes on the runoff generation of a river basin through parameter regionalization of a watershed model. *Journal of hydrology*, 292(1-4), 281–295.

Hundecha, Y., Ouarda, T. B. & Bárdossy, A. (2008). Regional estimation of parameters of a rainfall-runoff model at ungauged watersheds using the "spatial" structures of the parameters within a canonical physiographic-climatic space. *Water Resources Research*, 44(1).

Huot, P.-L. (2019). *Développement d'approches de calage efficaces pour les modèles hydrologiques coûteux en temps de calcul, au moyen du modèle HYDROTEL.* (Ph.D. thesis, École de technologie supérieure).

Huot, P.-L., Poulin, A., Audet, C. & Alarie, S. (2019). A hybrid optimization approach for efficient calibration of computationally intensive hydrological models. *Hydrological Sciences Journal*, (just-accepted).

Ichiba, A., Gires, A., Tchiguirinskaia, I., Schertzer, D., Bompard, P. & Veldhuis, M.-C. T. (2018). Scale effect challenges in urban hydrology highlighted with a distributed hydrological model. *Hydrology and Earth System Sciences*, 22(1), 331–350. Retrieved from: https://doi.org/10.5194/hess-22-331-2018.

IPCC, A. (2014). IPCC Fifth Assessment Report—Synthesis Report. IPCC New York, NY, USA.

Ishak, E. & Rahman, A. (2019). Examination of changes in flood data in Australia. *Water*, 11(8), 1734.

Ivancic, T. J. & Shaw, S. B. (2016). A US-based analysis of the ability of the Clausius-Clapeyron relationship to explain changes in extreme rainfall with changing temperature. *Journal of Geophysical Research: Atmospheres*, 121(7), 3066–3078.

James, G., Witten, D., Hastie, T. & Tibshirani, R. (2013). *An introduction to statistical learning*. Springer.

Javelle, P., Ouarda, T. B. & Bobée, B. (2003). Spring flood analysis using the flood-duration–frequency approach: application to the provinces of Quebec and Ontario, Canada. *Hydrological Processes*, 17(18), 3717–3736.

Jencso, K. G., McGlynn, B. L., Gooseff, M. N., Bencala, K. E. & Wondzell, S. M. (2010). Hillslope hydrologic connectivity controls riparian groundwater turnover: Implications of catchment structure for riparian buffering and stream water sources. *Water Resources Research*, 46(10).

Jeong, D. I. & Sushama, L. (2018). Rain-on-snow events over North America based on two Canadian regional climate models. *Climate dynamics*, 50(1-2), 303–316.

Jiang, D. & Wang, K. (2019). The role of satellite-based remote sensing in improving simulated streamflow: A review. *Water*, 11(8), 1615.

Jones, J. A., Swanson, F. J., Wemple, B. C. & Snyder, K. U. (2000). Effects of roads on hydrology, geomorphology, and disturbance patches in stream networks. *Conservation Biology*, 14(1), 76–85.

Joseph, J., Ghosh, S., Pathak, A. & Sahai, A. (2018). Hydrologic impacts of climate change: Comparisons between hydrological parameter uncertainty and climate model uncertainty. *Journal of Hydrology*, 566, 1–22. Retrieved from: https://doi.org/10.1016/j.jhydrol.2018.08.080.

Joyce, R. J., Janowiak, J. E., Arkin, P. A. & Xie, P. (2004). CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *Journal of Hydrometeorology*, 5(3), 487–503.

Jury, W. A. & Horton, R. (2004). *Soil physics*. John Wiley & Sons.

Kanishka, G. & Eldho, T. (2017). Watershed classification using isomap technique and hydrometeorological attributes. *Journal of Hydrologic Engineering*, 22(10), 04017040.

Kanishka, G. & Eldho, T. (2020). Streamflow estimation in ungauged basins using watershed classification and regionalization techniques. *Journal of Earth System Science*, 129(1), 1–18.

Kasiviswanathan, K. S., He, J., Sudheer, K. & Tay, J.-H. (2016). Potential application of wavelet neural network ensemble to forecast streamflow for flood management. *Journal of hydrology*, 536, 161–173.

Kattelmann, R. & Dozier, J. (1999). Observations of snowpack ripening in the Sierra Nevada, California, USA. *Journal of Glaciology*, 45(151), 409–416.

Kelleners, T., Chandler, D. G., McNamara, J. P., Gribb, M. M. & Seyfried, M. (2009). Modeling the water and energy balance of vegetated areas with snow accumulation. *Vadose Zone Journal*, 8(4), 1013–1030.

Keller, T., Sandin, M., Colombi, T., Horn, R. & Or, D. (2019). Historical increase in agricultural machinery weights enhanced soil stress levels and adversely affected soil functioning. *Soil and Tillage Research*, 194, 104293.

Kemter, M., Merz, B., Marwan, N., Vorogushyn, S. & Blöschl, G. (2020). Joint trends in flood magnitudes and spatial extents across Europe. *Geophysical Research Letters*, 47(7), e2020GL087464.

Kendon, E. J., Ban, N., Roberts, N. M., Fowler, H. J., Roberts, M. J., Chan, S. C., Evans, J. P., Fosser, G. & Wilkinson, J. M. (2017). Do convection-permitting regional climate models improve projections of future precipitation change? *Bulletin of the American Meteorological Society*, 98(1), 79–93.

Kharin, V. V., Zwiers, F. W., Zhang, X. & Hegerl, G. C. (2007). Changes in temperature and precipitation extremes in the IPCC ensemble of global coupled model simulations. *Journal of Climate*, 20(8), 1419–1444.

Kharin, V. V., Zwiers, F., Zhang, X. & Wehner, M. (2013). Changes in temperature and precipitation extremes in the CMIP5 ensemble. *Climatic change*, 119(2), 345–357.

Kharin, V., Flato, G., Zhang, X., Gillett, N., Zwiers, F. & Anderson, K. (2018). Risks from climate extremes change differently from 1.5 C to 2.0 C depending on rarity. *Earth's Future*, 6(5), 704–715. Retrieved from: https://doi.org/10.1002/2018EF000813.

Kim, J., Johnson, L., Cifelli, R., Thorstensen, A. & Chandrasekar, V. (2019). Assessment of antecedent moisture condition on flood frequency: An experimental study in Napa River Basin, CA. *Journal of Hydrology: Regional Studies*, 26, 100629. Retrieved from: https://doi.org/10.1016/j.ejrh.2019.100629.

Kim, U. & Kaluarachchi, J. J. (2008). Application of parameter estimation and regionalization methodologies to ungauged basins of the Upper Blue Nile River Basin, Ethiopia. *Journal of Hydrology*, 362(1-2), 39–56.

King, A. D., Donat, M. G., Fischer, E. M., Hawkins, E., Alexander, L. V., Karoly, D. J., Dittus, A. J., Lewis, S. C. & Perkins, S. E. (2015). The timing of anthropogenic emergence in simulated climate extremes. *Environmental Research Letters*, 10(9), 094015.

Kirchner, J. W. (2006). Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology. *Water Resources Research*, 42(3).

Kirpich, Z. (1940). Time of concentration of small agricultural watersheds. *Civil engineering*, 10(6), 362.

Kitanidis, P. K. & VoMvoris, E. G. (1983). A geostatistical approach to the inverse problem in groundwater modeling (steady state) and one-dimensional simulations. *Water Resources Research*, 19(3), 677–690. Retrieved from: https://doi.org/10.1029/WR019i003p00677.

Knutti, R., Furrer, R., Tebaldi, C., Cermak, J. & Meehl, G. A. (2010). Challenges in combining projections from multiple climate models. *Journal of Climate*, 23(10), 2739–2758.

Koci, J., Sidle, R. C., Jarihani, B. & Cashman, M. J. (2020). Linking hydrological connectivity to gully erosion in savanna rangelands tributary to the Great Barrier Reef using structure-from-motion photogrammetry. *Land Degradation & Development*, 31(1), 20–36.

Kratzert, F., Klotz, D., Brenner, C., Schulz, K. & Herrnegger, M. (2018). Rainfall–runoff modelling using long short-term memory (LSTM) networks. *Hydrology and Earth System Sciences*, 22(11), 6005–6022.

Kratzert, F., Klotz, D., Herrnegger, M., Sampson, A. K., Hochreiter, S. & Nearing, G. S. (2019a). Toward improved predictions in ungauged basins: Exploiting the power of machine learning. *Water Resources Research*, 55(12), 11344–11354.

Kratzert, F., Klotz, D., Shalev, G., Klambauer, G., Hochreiter, S. & Nearing, G. (2019b). Towards learning universal, regional, and local hydrological behaviors via machine learning applied to large-sample datasets. *Hydrology and Earth System Sciences*, 23(12), 5089–5110.

Kratzert, F., Klotz, D., Hochreiter, S. & Nearing, G. S. (2020). A note on leveraging synergy in multiple meteorological datasets with deep learning for rainfall-runoff modeling.

Krebs, G., Kokkonen, T., Valtanen, M., Setälä, H. & Koivusalo, H. (2014). Spatial resolution considerations for urban hydrological modelling. *Journal of Hydrology*, 512, 482–497. Retrieved from: https://doi.org/10.1016/j.jhydrol.2014.03.013.

Kreibich, H., Di Baldassarre, G., Vorogushyn, S., Aerts, J. C., Apel, H., Aronica, G. T., Arnbjerg-Nielsen, K., Bouwer, L. M., Bubeck, P., Caloiero, T. et al. (2017). Adaptation to flood risk: Results of international paired flood event studies. *Earth's Future*, 5(10), 953–965.

Krueger, T., Freer, J., Quinton, J. N., Macleod, C. J., Bilotta, G. S., Brazier, R. E., Butler, P. & Haygarth, P. M. (2010). Ensemble evaluation of hydrological model hypotheses. *Water Resources Research*, 46(7).

Krysanova, V., Vetter, T., Eisner, S., Huang, S., Pechlivanidis, I., Strauch, M., Gelfan, A., Kumar, R., Aich, V., Arheimer, B. et al. (2017). Intercomparison of regional-scale hydrological models and climate change impacts projected for 12 large river basins worldwide—a synthesis. *Environmental Research Letters*, 12(10), 105002.

Krysanova, V., Donnelly, C., Gelfan, A., Gerten, D., Arheimer, B., Hattermann, F. & Kundzewicz, Z. W. (2018). How the performance of hydrological models relates to credibility of projections under climate change. *Hydrological sciences journal*, 63(5), 696–720.

Krzysztofowicz, R. (2001). The case for probabilistic forecasting in hydrology. *Journal of hydrology*, 249(1-4), 2–9. Retrieved from: https://doi.org/10.1016/S0022-1694(01)00420-6.

Kuczera, G., Lambert, M., Heneker, T., Jennings, S., Frost, A. & Coombes, P. (2006). Joint probability and design storms at the crossroads. *Australasian Journal of Water Resources*, 10(1), 63–79.

Kumar, R., Samaniego, L. & Attinger, S. (2010). The effects of spatial discretization and model parameterization on the prediction of extreme runoff characteristics. *Journal of Hydrology*, 392(1-2), 54–69. Retrieved from: https://doi.org/10.1016/j.jhydrol.2010.07.047.

Kumar, R., Samaniego, L. & Attinger, S. (2013). Implications of distributed hydrologic model parameterization on water fluxes at multiple scales and locations. *Water Resources Research*, 49(1), 360–379. Retrieved from: https://doi.org/10.1029/2012WR012195.

Kundzewicz, Z. W. & Jun, X. (2004). Towards an improved flood preparedness system in China. Taylor & Francis.

Kundzewicz, Z. W. & Stakhiv, E. Z. (2010). Are climate models "ready for prime time" in water resources management applications, or is more research needed? *Hydrological Sciences Journal–Journal des Sciences Hydrologiques*, 55(7), 1085–1089.

Kundzewicz, Z. W., Graczyk, D., Maurer, T., Pińskwar, I., Radziejewski, M., Svensson, C. & Szwed, M. (2005). Trend detection in river flow series: 1. Annual maximum flow/Détection de tendance dans des séries de débit fluvial: 1. Débit maximum annuel. *Hydrological Sciences Journal*, 50(5).

Kundzewicz, Z. W., Kanae, S., Seneviratne, S. I., Handmer, J., Nicholls, N., Peduzzi, P., Mechler, R., Bouwer, L. M., Arnell, N., Mach, K. et al. (2014). Flood risk and climate change: global and regional perspectives. *Hydrological Sciences Journal*, 59(1), 1–28.

Kundzewicz, Z. W., Su, B., Wang, Y., Wang, G., Wang, G., Huang, J. & Jiang, T. (2019a). Flood risk in a range of spatial perspectives–from global to local scales. *Natural Hazards and Earth System Sciences*, 19(7), 1319–1328.

Kundzewicz, Z. W., Szwed, M. & Pińskwar, I. (2019b). Climate variability and floods—A global review. *Water*, 11(7), 1399.

Kundzewicz, Z., Krysanova, V., Benestad, R., Hov, Ø., Piniewski, M. & Otto, I. (2018). Uncertainty in climate change impacts on water resources. *Environmental Science & Policy*, 79, 1–8.

Leduc, M., Mailhot, A., Frigon, A., Martel, J.-L., Ludwig, R., Brietzke, G. B., Giguère, M., Brissette, F., Turcotte, R., Braun, M. et al. (2019). The ClimEx Project: a 50-member ensemble of climate change projections at 12-km resolution over Europe and Northeastern North America with the Canadian regional climate model (CRCM5). *Journal of Applied Meteorology and Climatology*, 58(4), 663–693.

Lenderink, G. & Van Meijgaard, E. (2010). Linking increases in hourly precipitation extremes to atmospheric temperature and moisture changes. *Environmental Research Letters*, 5(2), 025208.

Leonard, M., Westra, S., Phatak, A., Lambert, M., van den Hurk, B., McInnes, K., Risbey, J., Schuster, S., Jakob, D. & Stafford-Smith, M. (2014). A compound event framework for understanding extreme impacts. *Wiley Interdisciplinary Reviews: Climate Change*, 5(1), 113–128.

Li, Q., Zhu, Y., Shangguan, W., Wang, X., Li, L. & Yu, F. (2022). An attention-aware LSTM model for soil moisture and soil temperature prediction. *Geoderma*, 409, 115651.

Li, Y., Chang, J., Luo, L., Wang, Y., Guo, A., Ma, F. & Fan, J. (2019). Spatiotemporal impacts of land use land cover changes on hydrology from the mechanism perspective using SWAT model with time-varying parameters. *Hydrology Research*, 50(1), 244–261. Retrieved from: https://doi.org/10.2166/nh.2018.006.

Li, Z., Shao, Q., Xu, Z. & Cai, X. (2010). Analysis of parameter uncertainty in semi-distributed hydrological models using bootstrap method: A case study of SWAT model applied to Yingluoxia watershed in northwest China. *Journal of Hydrology*, 385(1-4), 76–83. Retrieved from: https://doi.org/10.1016/j.jhydrol.2010.01.025.

Liang, X., Lettenmaier, D. P., Wood, E. F. & Burges, S. J. (1994). A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *Journal of Geophysical Research: Atmospheres*, 99(D7), 14415–14428. Retrieved from: https://doi.org/10.1029/94JD00483.

Lindström, G., Johansson, B., Persson, M., Gardelin, M. & Bergström, S. (1997). Development and test of the distributed HBV-96 hydrological model. *Journal of hydrology*, 201(1-4), 272–288.

Lins, H. F. & Slack, J. R. (1999). Streamflow trends in the United States. *Geophysical research letters*, 26(2), 227–230.

Lins, H. F. & Slack, J. R. (2005). Seasonal and regional characteristics of US streamflow trends in the United States from 1940 to 1999. *Physical Geography*, 26(6), 489–501.

Lochbihler, K., Lenderink, G. & Siebesma, A. P. (2017). The spatial extent of rainfall events and its relation to precipitation scaling. *Geophysical Research Letters*, 44(16), 8629–8636.

López-Vicente, M., Sun, X., Onda, Y., Kato, H., Gomi, T. & Hiraoka, M. (2017). Effect of tree thinning and skidding trails on hydrological connectivity in two Japanese forest catchments. *Geomorphology*, 292, 104–114.

Lorenz, E. N. (1969). The predictability of a flow which possesses many scales of motion. *Tellus*, 21(3), 289–307. Retrieved from: https://doi.org/10.3402/tellusa.v21i3.10086.

Lorenzelli, F. (2014). *The essence of chaos*. CRC Press.

Lucas-Picher, P., Lachance-Cloutier, S., Arsenault, R., Poulin, A., Ricard, S., Turcotte, R. & Brissette, F. (2020). Will Evolving Climate Conditions Increase the Risk of Floods of the Large US-Canada Transboundary Richelieu River Basin? *JAWRA Journal of the American Water Resources Association*. Retrieved from: https://doi.org/10.1111/1752-1688.12891.

Lucas-Picher, P., Argüeso, D., Brisson, E., Tramblay, Y., Berg, P., Lemonsu, A., Kotlarski, S. & Caillaud, C. (2021). Convection-permitting modeling with regional climate models: Latest developments and next steps. *Wiley Interdisciplinary Reviews: Climate Change*, 12(6), e731.

Madsen, H., Lawrence, D., Lang, M., Martinkova, M. & Kjeldsen, T. (2014). Review of trend analysis and climate change projections of extreme precipitation and floods in Europe. *Journal of Hydrology*, 519, 3634–3650.

Mai, J., Craig, J. R., Tolson, B. A. & Arsenault, R. (2022a). The sensitivity of simulated stream-flow to individual hydrologic processes across North America. *Nature communications*, 13(1), 1–11.

Mai, J., Shen, H., Tolson, B. A., Gaborit, É., Arsenault, R., Craig, J. R., Fortin, V., Fry, L. M., Gauch, M., Klotz, D. et al. (2022b). The Great Lakes Runoff Intercomparison Project Phase 4: the Great Lakes (GRIP-GL). *Hydrology and Earth System Sciences*, 26(13), 3537–3572.

Maraun, D. (2013). When will trends in European mean and heavy daily precipitation emerge? *Environmental Research Letters*, 8(1), 014004.

Maraun, D. & Widmann, M. (2018). *Statistical downscaling and bias correction for climate research*. Cambridge University Press.

Maraun, D., Wetterhall, F., Ireson, A., Chandler, R., Kendon, E., Widmann, M., Brienen, S., Rust, H., Sauter, T., Themeßl, M. et al. (2010). Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user. *Reviews of Geophysics*, 48(3).

Markhali, S. P., Poulin, A. & Boucher, M.-A. (2022). Spatio-Temporal Discretization Uncertainty of Distributed Hydrological Models. *Hydrological Processes*, e14635.

Marks, D. & Dozier, J. (1992). Climate and energy exchange at the snow surface in the alpine region of the Sierra Nevada: 2. Snow cover energy balance. *Water Resources Research*, 28(11), 3043–3054.

Martel, J.-L. (2019). *Évaluation de l'influence de la variabilité naturelle du climat et des changements climatiques anthropiques sur les extrêmes hydrométéorologiques*. (Ph.D. thesis, École de technologie supérieure).

Martel, J.-L., Mailhot, A., Brissette, F. & Caya, D. (2018). Role of natural climate variability in the detection of anthropogenic climate change signal for mean and extreme precipitation at local and regional scales. *Journal of Climate*, 31(11), 4241–4263.

Martel, J.-L., Brissette, F. & Poulin, A. (2020a). Impact of the spatial density of weather stations on the performance of distributed and lumped hydrological models. *Canadian Water Resources Journal/Revue canadienne des ressources hydriques*, 1–14. Retrieved from: https://doi.org/10.5194/gmd-12-2501-2019.

Martel, J.-L., Mailhot, A. & Brissette, F. (2020b). Global and regional projected changes in 100-yr subdaily, daily, and multiday precipitation extremes estimated from three large ensembles of climate simulations. *Journal of Climate*, 33(3), 1089–1103.

Martel, J.-L., Brissette, F. P., Lucas-Picher, P., Troin, M. & Arsenault, R. (2021). Climate change and rainfall intensity-duration-frequency curves: Overview of science and guidelines for adaptation. *Journal of Hydrologic Engineering*, 26(10).

Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M. et al. (2021). Climate change 2021: the physical science basis. *Contribution of working group I to the sixth assessment report of the intergovernmental panel on climate change*, 2.

Mazurkiewicz, A. B., Callery, D. G. & McDonnell, J. J. (2008). Assessing the controls of the snow energy balance and water available for runoff in a rain-on-snow environment. *Journal of Hydrology*, 354(1-4), 1–14.

McCabe, G. J., Clark, M. P. & Hay, L. E. (2007). Rain-on-snow events in the western United States. *Bulletin of the American Meteorological Society*, 88(3), 319–328.

McDonnell, J., Sivapalan, M., Vaché, K., Dunn, S., Grant, G., Haggerty, R., Hinz, C., Hooper, R., Kirchner, J., Roderick, M. et al. (2007). Moving beyond heterogeneity and process complexity: A new vision for watershed hydrology. *Water Resources Research*, 43(7).

McGlynn, B. L. & McDonnell, J. J. (2003). Quantifying the relative contributions of riparian and hillslope zones to catchment runoff. *Water Resources Research*, 39(11).

McInerney, D., Thyer, M., Kavetski, D., Lerat, J. & Kuczera, G. (2017). Improving probabilistic prediction of daily streamflow by identifying P areto optimal approaches for modeling heteroscedastic residual errors. *Water Resources Research*, 53(3), 2199–2239.

McMillan, H. K., Westerberg, I. K. & Krueger, T. (2018). Hydrological data uncertainty and its implications. *Wiley Interdisciplinary Reviews: Water*, 5(6), e1319.

Mearns, L., McGinnis, S., Korytina, D., Arritt, R., Biner, S., Bukovsky, M., Chang, H., Christensen, O., Herzmann, D., Jiao, Y. et al. (2017). The NA-CORDEX dataset, version 1.0. NCAR climate data gateway, boulder CO.

Meehl, G. A., Covey, C., Delworth, T., Latif, M., McAvaney, B., Mitchell, J. F., Stouffer, R. J. & Taylor, K. E. (2007). The WCRP CMIP3 multimodel dataset: A new era in climate change research. *Bulletin of the American Meteorological Society*, 88(9), 1383–1394.

Mei, X., Van Gelder, P., Dai, Z. & Tang, Z. (2017). Impact of dams on flood occurrence of selected rivers in the United States. *Frontiers of Earth Science*, 11(2), 268–282.

Meinshausen, M., Smith, S. J., Calvin, K., Daniel, J. S., Kainuma, M., Lamarque, J.-F., Matsumoto, K., Montzka, S., Raper, S., Riahi, K. et al. (2011). The RCP greenhouse gas concentrations and their extensions from 1765 to 2300. *Climatic change*, 109(1-2), 213.

Meresa, H., Tischbein, B., Mendela, J., Demoz, R., Abreha, T., Weldemichael, M. & Ogbu, K. (2022). The role of input and hydrological parameters uncertainties in extreme hydrological simulations. *Natural Resource Modeling*, 35(1), e12320.

Meresa, H. K. & Romanowicz, R. J. (2017). The critical role of uncertainty in projections of hydrological extremes. *Hydrology and Earth System Sciences*, 21(8), 4245–4258.

Merz, B., Blöschl, G., Vorogushyn, S., Dottori, F., Aerts, J. C., Bates, P., Bertola, M., Kemter, M., Kreibich, H., Lall, U. et al. (2021). Causes, impacts and patterns of disastrous river floods. *Nature Reviews Earth & Environment*, 2(9), 592–609.

Merz, R. & Blöschl, G. (2003). A process typology of regional floods. *Water Resources Research*, 39(12).

Merz, R. & Blöschl, G. (2004). Regionalisation of catchment model parameters. *Journal of hydrology*, 287(1-4), 95–123. Retrieved from: https://doi.org/10.1016/j.jhydrol.2003.09.028.

Merz, R. & Blöschl, G. (2009). A regional analysis of event runoff coefficients with respect to climate and catchment characteristics in Austria. *Water Resources Research*, 45(1).

Merz, R., Tarasova, L. & Basso, S. (2020). Parameter's controls of distributed catchment models—How much information is in conventional catchment descriptors? *Water Resources Research*, 56(2), e2019WR026008.

Meyer, J., Kohn, I., Stahl, K., Hakala, K., Seibert, J. & Cannon, A. J. (2019). Effects of univariate and multivariate bias correction on hydrological impact projections in alpine catchments. *Hydrology and Earth System Sciences*, 23(3), 1339–1354.

Miller, J. D. & Hutchins, M. (2017). The impacts of urbanisation and climate change on urban flooding and urban water quality: A review of the evidence concerning the United Kingdom. *Journal of Hydrology: Regional Studies*, 12, 345–362.

Min, S.-K., Zhang, X., Zwiers, F. W. & Hegerl, G. C. (2011). Human contribution to more-intense precipitation extremes. *Nature*, 470(7334), 378.

Minville, M., Brissette, F. & Leconte, R. (2008). Uncertainty of the impact of climate change on the hydrology of a nordic watershed. *Journal of hydrology*, 358(1-2), 70–83.

Miyata, S., Gomi, T., Sidle, R. C., Hiraoka, M., Onda, Y., Yamamoto, K. & Nonoda, T. (2019). Assessing spatially distributed infiltration capacity to evaluate storm runoff in forested catchments: Implications for hydrological connectivity. *Science of the Total Environment*, 669, 148–159.

Moftakhari, H., Schubert, J. E., AghaKouchak, A., Matthew, R. A. & Sanders, B. F. (2019). Linking statistical and hydrodynamic modeling for compound flood hazard assessment in tidal channels and estuaries. *Advances in Water Resources*, 128, 28–38.

Moftakhari, H. R., AghaKouchak, A., Sanders, B. F. & Matthew, R. A. (2017). Cumulative hazard: The case of nuisance flooding. *Earth's Future*, 5(2), 214–223.

Molnar, P., Fatichi, S., Gaál, L., Szolgay, J. & Burlando, P. (2015). Storm type effects on super Clausius–Clapeyron scaling of intense rainstorm properties with air temperature. *Hydrology and Earth System Sciences*, 19(4), 1753–1766.

Montanari, A. & Koutsoyiannis, D. (2012). A blueprint for process-based modeling of uncertain hydrological systems. *Water Resources Research*, 48(9).

Montanari, A. & Koutsoyiannis, D. (2014). Modeling and mitigating natural hazards: Stationarity is immortal! *Water resources research*, 50(12), 9748–9756.

Montgomery, D. C. (2017). *Design and analysis of experiments*. John wiley & sons.

Mooney, P., Broderick, C., Bruyère, C., Mulligan, F. & Prein, A. (2017). Clustering of observed diurnal cycles of precipitation over the United States for evaluation of a WRF multiphysics regional climate ensemble. *Journal of Climate*, 30(22), 9267–9286.

Moore, I. D. & Grayson, R. B. (1991). Terrain-based catchment partitioning and runoff prediction using vector elevation data. *Water Resources Research*, 27(6), 1177–1191.

Moore, R. (1985). The probability-distributed principle and runoff production at point and basin scales. *Hydrological Sciences Journal*, 30(2), 273–297.

Muñoz, P., Orellana-Alvear, J., Willems, P. & Célleri, R. (2018). Flash-flood forecasting in an Andean mountain catchment—development of a step-wise methodology based on the random forest algorithm. *Water*, 10(11), 1519.

Nakamura, J., Lall, U., Kushnir, Y., Robertson, A. W. & Seager, R. (2013). Dynamical structure of extreme floods in the US Midwest and the United Kingdom. *Journal of Hydrometeorology*, 14(2), 485–504.

Nash, J. E. & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I—A discussion of principles. *Journal of hydrology*, 10(3), 282–290.

Nasseri, M., Zahraie, B. & Tootchi, A. (2019). Spatial scale resolution of prognostic hydrological models: simulation performance and application in climate change impact assessment. *Water Resources Management*, 33(1), 189–205.

Navarro-Racines, C., Tarapues, J., Thornton, P., Jarvis, A. & Ramirez-Villegas, J. (2020). High-resolution and bias-corrected CMIP5 projections for climate change impact assessments. *Scientific data*, 7(1), 1–14.

Nearing, G. & Gupta, H. (2018a). Philosophical Foundations of Hydrologic Uncertainty. *EGU General Assembly Conference Abstracts*, 20, 5369.

Nearing, G. S. & Gupta, H. V. (2015). The quantity and quality of information in hydrologic models. *Water Resources Research*, 51(1), 524–538. Retrieved from: https://doi.org/10.1002/2014WR015895.

Nearing, G. S. & Gupta, H. V. (2018b). Ensembles vs. information theory: supporting science under uncertainty. *Frontiers of Earth Science*, 12(4), 653–660. Retrieved from: https://doi.org/10.1007/s11707-018-0709-9.

Nearing, G. S., Gupta, H. V. & Crow, W. T. (2013). Information loss in approximately Bayesian estimation techniques: A comparison of generative and discriminative approaches to estimating agricultural productivity. *Journal of hydrology*, 507, 163–173. Retrieved from: https://doi.org/10.1016/j.jhydrol.2013.10.029.

Nearing, G. S., Mocko, D. M., Peters-Lidard, C. D., Kumar, S. V. & Xia, Y. (2016). Benchmarking NLDAS-2 soil moisture and evapotranspiration to separate uncertainty contributions. *Journal of hydrometeorology*, 17(3), 745–759.

Nearing, G. S., Kratzert, F., Sampson, A. K., Pelissier, C. S., Klotz, D., Frame, J. M., Prieto, C. & Gupta, H. V. (2020). What role does hydrological science play in the age of machine learning? *Water Resources Research*, e2020WR028091. Retrieved from: https://doi.org/10.1029/2020WR028091.

Nelson, R. (1999). An Introduction to Copulas.

Neuman, S. P. (1990). Universal scaling of hydraulic conductivities and dispersivities in geologic media. *Water resources research*, 26(8), 1749–1758. Retrieved from: https://doi.org/10.1029/WR026i008p01749.

Noguchi, S., Tsuboyama, Y., Sidle, R. C. & Hosoda, I. (1999). Morphological characteristics of macropores and the distribution of preferential flow pathways in a forested slope segment. *Soil Science Society of America Journal*, 63(5), 1413–1423.

Olden, J. D. & Poff, N. (2003). Redundancy and the choice of hydrologic indices for characterizing streamflow regimes. *River research and applications*, 19(2), 101–121.

O'Neill, B. C., Tebaldi, C., Vuuren, D. P. v., Eyring, V., Friedlingstein, P., Hurtt, G., Knutti, R., Kriegler, E., Lamarque, J.-F., Lowe, J. et al. (2016). The scenario model intercomparison project (ScenarioMIP) for CMIP6. *Geoscientific Model Development*, 9(9), 3461–3482.

Oubennaceur, K., Chokmani, K., Nastev, M., Lhissou, R. & El Alem, A. (2019). Flood risk mapping for direct damage to residential buildings in Quebec, Canada. *International Journal of Disaster Risk Reduction*, 33, 44–54.

Oyebode, O. & Stretch, D. (2019). Neural network modeling of hydrological systems: A review of implementation techniques. *Natural Resource Modeling*, 32(1), e12189.

O'Gorman, P. A. (2015). Precipitation extremes under climate change. *Current climate change reports*, 1(2), 49–59.

Pachauri, R. K., Allen, M. R., Barros, V. R., Broome, J., Cramer, W., Christ, R., Church, J. A., Clarke, L., Dahe, Q., Dasgupta, P. et al. (2014). *Climate change 2014: synthesis report. Contribution of Working Groups I, II and III to the fifth assessment report of the Intergovernmental Panel on Climate Change*. Ipcc.

Pall, P., Allen, M. & Stone, D. A. (2007). Testing the Clausius–Clapeyron constraint on changes in extreme precipitation under CO 2 warming. *Climate Dynamics*, 28(4), 351–363.

Pandey, B. K., Khare, D., Kawasaki, A. & Mishra, P. K. (2019). Climate change impact assessment on blue and green water by coupling of representative CMIP5 climate models with physical based hydrological model. *Water resources management*, 33(1), 141–158.

Papacharalampous, G., Tyralis, H. & Koutsoyiannis, D. (2019). Comparison of stochastic and machine learning methods for multi-step ahead forecasting of hydrological processes. *Stochastic Environmental Research and Risk Assessment*, 33(2), 481–514. Retrieved from: https://doi.org/10.1007/s00477-018-1638-6.

Papageorgaki, I. & Nalbantis, I. (2016). Classification of drainage basins based on readily available information. *Water Resources Management*, 30(15), 5559–5574.

Paprotny, D., Vousdoukas, M. I., Morales-Nápoles, O., Jonkman, S. N. & Feyen, L. (2018). Compound flood potential in Europe. *Hydrol. Earth Syst. Sci. Discuss*, 2018, 1–34.

Parajka, J., Viglione, A., Rogger, M., Salinas, J., Sivapalan, M. & Blöschl, G. (2013). Comparative assessment of predictions in ungauged basins–Part 1: Runoff-hydrograph studies. *Hydrology and Earth System Sciences*, 17(5), 1783–1795.

Parajka, J., Merz, R. & Blöschl, G. (2005). A comparison of regionalisation methods for catchment model parameters. *Hydrology and Earth System Sciences*, 9(3), 157–171.

Pathiraja, S., Westra, S. & Sharma, A. (2012). Why continuous simulation? The role of antecedent moisture in design flood estimation. *Water Resources Research*, 48(6). doi: http://dx.doi.org/10.1002/andp.19053221004.

Perrin, C., Michel, C. & Andréassian, V. (2003). Improvement of a parsimonious model for streamflow simulation. *Journal of hydrology*, 279(1-4), 275–289.

Peters-Lidard, C. D., Clark, M., Samaniego, L., Verhoest, N. E., Van Emmerik, T., Uijlenhoet, R., Achieng, K., Franz, T. E. & Woods, R. (2017). Scaling, similarity, and the fourth paradigm for hydrology. *Hydrology and earth system sciences*, 21(7), 3701. Retrieved from: https://doi.org/10.5194/hess-21-3701-2017.

Petvirojchai, P. & SaraPa, S. (2018). Current technology for alerting and warning tropical cyclones in Thailand. *Tropical Cyclone Research and Review*, 7(3), 193–200.

Pham, H. T., Marshall, L., Johnson, F. & Sharma, A. (2018). A method for combining SRTM DEM and ASTER GDEM2 to improve topography estimation in regions without reference data. *Remote sensing of environment*, 210, 229–241.

Phillips, T. J., Potter, G. L., Williamson, D. L., Cederwall, R. T., Boyle, J. S., Fiorino, M., Hnilo, J. J., Olson, J. G., Xie, S. & Yio, J. J. (2004). Evaluating parameterizations in general circulation models: Climate simulation meets weather prediction. *Bulletin of the American Meteorological Society*, 85(12), 1903–1916.

Pitie, F., Kokaram, A. C. & Dahyot, R. (2005). N-dimensional probability density function transfer and its application to color transfer. *Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1*, 2, 1434–1439.

Pokhrel, P., Gupta, H. & Wagener, T. (2008). A spatial regularization approach to parameter estimation for a distributed watershed model. *Water Resources Research*, 44(12).

Pokhrel, P. & Gupta, H. V. (2010). On the use of spatial regularization strategies to improve calibration of distributed watershed models. *Water Resources Research*, 46(1). Retrieved from: https://doi.org/10.1029/2009WR008066.

Pomeroy, J. W., Fang, X. & Marks, D. G. (2016). The cold rain-on-snow event of June 2013 in the Canadian Rockies—characteristics and diagnosis. *Hydrological Processes*, 30(17), 2899–2914.

Poulin, A., Brissette, F., Leconte, R., Arsenault, R. & Malo, J.-S. (2011). Uncertainty of hydrological modelling in climate change impact studies in a Canadian, snow-dominated river basin. *Journal of Hydrology*, 409(3-4), 626–636. Retrieved from: https://doi.org/10.1016/j.jhydrol.2011.08.057.

Prein, A. F., Langhans, W., Fosser, G., Ferrone, A., Ban, N., Goergen, K., Keller, M., Tölle, M., Gutjahr, O., Feser, F. et al. (2015). A review on regional convection-permitting climate modeling: Demonstrations, prospects, and challenges. *Reviews of geophysics*, 53(2), 323–361.

Prein, A. F., Rasmussen, R. M., Ikeda, K., Liu, C., Clark, M. P. & Holland, G. J. (2017). The future intensification of hourly precipitation extremes. *Nature Climate Change*, 7(1), 48–52.

Pringle, C. M. (2001). Hydrologic connectivity and the management of biological reserves: a global perspective. *Ecological Applications*, 11(4), 981–998.

Pui, A., Lal, A. & Sharma, A. (2011). How does the Interdecadal Pacific Oscillation affect design floods in Australia? *Water Resources Research*, 47(5).

Quilty, J., Adamowski, J. & Boucher, M.-A. (2019). A stochastic data-driven ensemble forecasting framework for water resources: A case study using ensemble members derived from a database of deterministic wavelet-based models. *Water Resources Research*, 55(1), 175–202. Retrieved from: https://doi.org/10.1029/2018WR023205.

Rajaram, H. (2016). Debates-Stochastic subsurface hydrology from theory to practice: Introduction. *Water Resources Research*, 52(12), 9215–9217.

Rakovec, O., Kumar, R., Mai, J., Cuntz, M., Thober, S., Zink, M., Attinger, S., Schäfer, D., Schrön, M. & Samaniego, L. (2016). Multiscale and multivariate evaluation of water fluxes and states over European river basins. *Journal of Hydrometeorology*, 17(1), 287–307. Retrieved from: https://doi.org/10.1175/JHM-D-15-0054.1.

Randrianasolo, A., Ramos, M. & Andréassian, V. (2011). Hydrological ensemble forecasting at ungauged basins: using neighbour catchments for model setup and updating. *Advances in Geosciences*, 29, 1–11.

Razavi, S., Tolson, B. A. & Burn, D. H. (2012). Review of surrogate modeling in water resources. *Water Resources Research*, 48(7).

Razavi, T. & Coulibaly, P. (2013). Streamflow prediction in ungauged basins: review of regionalization methods. *Journal of hydrologic engineering*, 18(8), 958–975.

Razavi, T. & Coulibaly, P. (2017). An evaluation of regionalization and watershed classification schemes for continuous daily streamflow prediction in ungauged watersheds. *Canadian Water Resources Journal/Revue canadienne des ressources hydriques*, 42(1), 2–20.

Reed, S., Koren, V., Smith, M., Zhang, Z., Moreda, F., Seo, D.-J. & Participants, D. (2004). Overall distributed model intercomparison project results. *Journal of Hydrology*, 298(1-4), 27–60.

Refsgaard, C. (1995). Mike she. *Computer models of catchment hydrology*, 809–846. Retrieved from: https://scholar.google.ca/scholar?hl=en&as_sdt=0%2C5&as_vis=1&q=Mike+she&btnG=.

Refsgaard, J. C. & Storm, B. (1990). Construction, calibration and validation of hydrological models. In *Distributed hydrological modelling* (pp. 41–54). Springer.

Refsgaard, J. C., Van der Sluijs, J. P., Brown, J. & Van der Keur, P. (2006). A framework for dealing with uncertainty due to model structure error. *Advances in Water Resources*, 29(11), 1586–1597. Retrieved from: https://doi.org/10.1016/j.advwatres.2005.11.013.

Resnick, S. (1987). Extreme values, point processes and regular variation. Springer Verlag, New York.

Riboust, P. & Brissette, F. (2016). Analysis of Lake Champlain/Richelieu River's historical 2011 flood. *Canadian Water Resources Journal/Revue canadienne des ressources hydriques*, 41(1-2), 174–185.

Richards, L. A. (1931). Capillary conduction of liquids through porous mediums. *Physics*, 1(5), 318–333.

Rigon, R., Formetta, G., Bancheri, M., Tubini, N., D'Amato, C., David, O. & Massari, C. (2022). *Participatory Digital Earth Twin Hydrology systems (DARTHs) for everyone: a blueprint for hydrologists*.

Rocheta, E., Evans, J. P. & Sharma, A. (2014). Assessing atmospheric bias correction for dynamical consistency using potential vorticity. *Environmental Research Letters*, 9(12), 124010.

Rogger, M., Agnoletti, M., Alaoui, A., Bathurst, J., Bodner, G., Borga, M., Chaplot, V., Gallart, F., Glatzel, G., Hall, J. et al. (2017). Land use change impacts on floods at the catchment scale: Challenges and opportunities for future research. *Water resources research*, 53(7), 5209–5219.

Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, 1(5), 206–215.

Saad, C., St-Hilaire, A., Gachon, P. & El Adlouni, S. (2016). The 2011 flood event in the Richelieu River basin: causes, assessment and damages. *Canadian Water Resources Journal/Revue canadienne des ressources hydriques*, 41(1-2), 129–138.

Saadi, M., Oudin, L. & Ribstein, P. (2019). Random forest ability in regionalizing hourly hydrological model parameters. *Water*, 11(8), 1540.

Sabarly, F., Essou, G., Lucas-Picher, P., Poulin, A. & Brissette, F. (2016). Use of four reanalysis datasets to assess the terrestrial branch of the water cycle over Quebec, Canada. *Journal of Hydrometeorology*, 17(5), 1447–1466.

Sadegh, M., Moftakhari, H., Gupta, H. V., Ragno, E., Mazdiyasni, O., Sanders, B., Matthew, R. & AghaKouchak, A. (2018). Multihazard scenarios for analysis of compound extreme events. *Geophysical Research Letters*, 45(11), 5470–5480. doi: https://doi.org/10.1029/2018GL077317.

Samaniego, L., Kumar, R. & Attinger, S. (2010). Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale. *Water Resources Research*, 46(5). Retrieved from: https://doi.org/10.1029/2008WR007327.

Samaniego, L., Kumar, R., Thober, S., Rakovec, O., Zink, M., Wanders, N., Eisner, S., Müller Schmied, H., Sutanudjaja, E., Warrach-Sagi, K. et al. (2017). Toward seamless hydrologic predictions across spatial scales. *Hydrology and Earth System Sciences*, 21(9), 4323–4346. Retrieved from: https://doi.org/10.5194/hess-21-4323-2017.

Samuel, J., Coulibaly, P. & Metcalfe, R. A. (2011). Estimation of continuous streamflow in Ontario ungauged basins: comparison of regionalization methods. *Journal of Hydrologic Engineering*, 16(5), 447–459.

Sanderson, B. M., Oleson, K. W., Strand, W. G., Lehner, F. & O'Neill, B. C. (2018). A new ensemble of GCM simulations to assess avoided impacts in a climate mitigation scenario. *Climatic Change*, 146(3-4), 303–318.

Savenije, H. H. (2009). HESS Opinions" The art of hydrology". *Hydrology and Earth System Sciences*, 13(2), 157–161.

Savenije, H. H. & Hrachowitz, M. (2017). HESS Opinions Catchments as meta-organisms–a new blueprint for hydrological modelling. *Hydrology and Earth System Sciences*, 21(2), 1107–1116.

Schaefli, B. & Gupta, H. V. (2007). Do Nash values have value? *Hydrological Processes*, 21(ARTICLE), 2075–2080.

Schär, C., Fuhrer, O., Arteaga, A., Ban, N., Charpilloz, C., Di Girolamo, S., Hentgen, L., Hoefler, T., Lapillonne, X., Leutwyler, D. et al. (2020). Kilometer-scale climate models: Prospects and challenges. *Bulletin of the American Meteorological Society*, 101(5), E567–E587.

Schröter, K., Kunz, M., Elmer, F., Mühr, B. & Merz, B. (2015). What made the June 2013 flood in Germany an exceptional event? A hydro-meteorological evaluation. *Hydrology and Earth System Sciences*, 19(1), 309–327.

Schulla, J. & Jasper, K. (1998). Modellbeschreibung WaSiM-ETH. *ETH Zürich*.

Schulla, J. & Jasper, K. (2007). Model description wasim-eth. *Institute for Atmospheric and Climate Science, Swiss Federal Institute of Technology, Zürich*. Retrieved from: https://scholar.google.ca/scholar?hl=en&as_sdt=0%2C5&as_vis=1&q=Model+description+wasim-eth&btnG=.

Schwierz, C., Davies, H. C., Appenzeller, C., Liniger, M. A., Müller, W., Stocker, T. F. & Yoshimori, M. (2006). Challenges posed by and approaches to the study of seasonal-to-decadal climate variability. In *Climate Variability, Predictability and Climate Risks* (pp. 31–63). Springer.

Seager, R. & Vecchi, G. A. (2010). Greenhouse warming and the 21st century hydroclimate of southwestern North America. *Proceedings of the National Academy of Sciences*, 107(50), 21277–21282.

Sedell, J. R., Reeves, G. H., Hauer, F. R., Stanford, J. A. & Hawkins, C. P. (1990). Role of refugia in recovery from disturbances: modern fragmented and disconnected river systems. *Environmental management*, 14(5), 711–724.

Seiller, G., Anctil, F. & Roy, R. (2017). Design and experimentation of an empirical multistructure framework for accurate, sharp and reliable hydrological ensembles. *Journal of Hydrology*, 552, 313–340.

Setegn, S. G., Srinivasan, R. & Dargahi, B. (2008). Hydrological modelling in the Lake Tana Basin, Ethiopia using SWAT model. *The Open Hydrology Journal*, 2(1).

Shen, C. (2018). A transdisciplinary review of deep learning research and its relevance for water resources scientists. *Water Resources Research*, 54(11), 8558–8593.

Shen, C., Chen, X. & Laloy, E. (2021). Broadening the use of machine learning in hydrology. Frontiers Media SA.

Sidle, R. C. (2006). Field observations and process understanding in hydrology: essential components in scaling. *Hydrological Processes*, 20(6), 1439–1445.

Sidle, R. C. (2021). Strategies for smarter catchment hydrology models: incorporating scaling and better process representation. *Geoscience Letters*, 8(1), 1–14.

Sidle, R. C., Gomi, T., Usuga, J. C. L. & Jarihani, B. (2017). Hydrogeomorphic processes and scaling issues in the continuum from soil pedons to catchments. *Earth-Science Reviews*, 175, 75–96.

Sikorska, A. E., Viviroli, D. & Seibert, J. (2018). Effective precipitation duration for runoff peaks based on catchment modelling. *Journal of hydrology*, 556, 510–522.

Silasari, R., Parajka, J., Ressl, C., Strauss, P. & Blöschl, G. (2017). Potential of time-lapse photography for identifying saturation area dynamics on agricultural hillslopes. *Hydrological Processes*, 31(21), 3610–3627.

Šimunek, J., van Genuchten, M. T. & Šejna, M. (2008). Development and applications of the HYDRUS and STANMOD software packages and related codes. *Vadose Zone Journal*, 7(2), 587–600. Retrieved from: https://doi.org/10.2136/vzj2007.0077.

Singh, H. V., Kalin, L., Morrison, A., Srivastava, P., Lockaby, G. & Pan, S. (2015). Post-validation of SWAT model in a coastal watershed for predicting land use/cover change impacts. *Hydrology Research*, 46(6), 837–853. Retrieved from: https://doi.org/10.2166/nh.2015.222.

Singh, M. & Sinha, R. (2021). Hydrogeomorphic indicators of wetland health inferred from multi-temporal remote sensing data for a new Ramsar site (Kaabar Tal), India. *Ecological Indicators*, 127, 107739.

Singh, P., Spitzbart, G., Hübl, H. & Weinmeister, H. (1997). Hydrological response of snowpack under rain-on-snow events: a field study. *Journal of Hydrology*, 202(1-4), 1–20.

Singh, V. P. & Frevert, D. K. (2005). *Watershed models*. CRC press.

Sivapalan, M. (2003). Process complexity at hillslope scale, process simplicity at watershed scale: Is there a connection? *EGS-AGU-EUG Joint Assembly*, pp. 7973.

Sivapalan, M. & Blöschl, G. (2015). Time scale interactions and the coevolution of humans and water. *Water Resources Research*, 51(9), 6988–7022.

Sivapalan, M. & Blöschl, G. (2017). The growth of hydrological understanding: Technologies, ideas, and societal needs shape the field. *Water Resources Research*, 53(10), 8137–8146.

Sivapalan, M., Blöschl, G., Merz, R. & Gutknecht, D. (2005). Linking flood frequency to long-term water balance: Incorporating effects of seasonality. *Water Resources Research*, 41(6).

Snelder, T. H., Lamouroux, N., Leathwick, J. R., Pella, H., Sauquet, E. & Shankar, U. (2009). Predictive mapping of the natural flow regimes of France. *Journal of Hydrology*, 373(1-2), 57–67.

Solman, S. A. & Blázquez, J. (2019). Multiscale precipitation variability over South America: analysis of the added value of CORDEX RCM simulations. *Climate Dynamics*, 53(3), 1547–1565.

St-Hilaire, A., Boucher, M., Chebana, F., Ouellet-Proulx, S., Zhou, Q., Larabi, S. & Dugdale, S. (2015a). Breathing a new life to an older model: The CEQUEAU tool for flow and water temperature simulations and forecasting. *L'eau Pour Le Développement Durable: Adaptation Aux Changements Du Climat et de L'environnement*, 48(6). Retrieved from: https://scholar.google.ca/scholar?hl=en&as_sdt=0%2C5&as_vis=1&q=Breathing+a+new+life+to+an+older+model%3A+The+CEQUEAU+tool+for+flow+and+water+temperature+simulations+and+forecasting&btnG=.

St-Hilaire, A., Boucher, M.-A., Chebana, F., Ouellet-Proulx, S., Zhou, Q. X., Larabi, S., Dugdale, S. & Latraverse, M. (2015b). Breathing a new life to an older model: The CEQUEAU tool for flow and water temperature simulations and forecasting. *Proceedings of the 22nd Canadian Hydrotechnical Conference*.

Stafford-Smith, M. (2014). A compound event framework for understanding extreme impacts. *Wiley Interdiscip Rev Clim Chang*, 5, 113128McInnes.

Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., Midgley, P. M. et al. (2013). Climate change 2013: The physical science basis. Cambridge University Press Cambridge.

Stommel, H. (1963). Varieties of oceanographic experience. *Science*, 139(3555), 572–576.

Sultana, Z., Sieg, T., Kellermann, P., Müller, M. & Kreibich, H. (2018). Assessment of business interruption of flood-affected companies using random forests. *Water*, 10(8), 1049.

Swain, D., Wing, O. E., Bates, P. D., Done, J., Johnson, K. & Cameron, D. (2020). Increased flood exposure due to climate change and population growth in the United States. *Earth's Future*, 8(11), e2020EF001778.

Swart, N. C., Cole, J. N., Kharin, V. V., Lazare, M., Scinocca, J. F., Gillett, N. P., Anstey, J., Arora, V., Christian, J. R., Hanna, S. et al. (2019). The Canadian earth system model version 5 (CanESM5. 0.3). *Geoscientific Model Development*, 12(11), 4823–4873.

Szolgayova, E., Laaha, G., Blöschl, G. & Bucher, C. (2014). Factors influencing long range dependence in streamflow of European rivers. *Hydrological Processes*, 28(4), 1573–1586.

Tarek, M., Brissette, F. & Arsenault, R. (2020a). Uncertainty of gridded precipitation and temperature reference datasets in climate change impact studies. *Hydrology and Earth System Sciences Discussions*, 1–32. Retrieved from: https://doi.org/10.5194/hess-25-3331-2021.

Tarek, M., Brissette, F. P. & Arsenault, R. (2020b). Evaluation of the ERA5 reanalysis as a potential reference dataset for hydrological modelling over North America. *Hydrology and Earth System Sciences*, 24(5), 2527–2544. Retrieved from: https://doi.org/10.5194/hess-24-2527-2020.

Tavangar, S., Moradi, H., Massah Bavani, A. & Gholamalifard, M. (2019). A futuristic survey of the effects of LU/LC change on stream flow by CA–Markov model: a case of the Nekarood watershed, Iran. *Geocarto International*, 1–17. Retrieved from: https://doi.org/10.1080/10106049.2019.1633419.

Taylor, K. E., Stouffer, R. J. & Meehl, G. A. (2012). An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, 93(4), 485–498.

Te Linde, A., Aerts, J. & Kwadijk, J. (2010). Effectiveness of flood management measures on peak discharges in the Rhine basin under climate change. *Journal of Flood Risk Management*, 3(4), 248–269.

Tegegne, G., Kim, Y.-O., Seo, S. B. & Kim, Y. (2019). Hydrological modelling uncertainty analysis for different flow quantiles: a case study in two hydro-geographically different watersheds. *Hydrological Sciences Journal*, 64(4), 473–489. Retrieved from: https://doi.org/10.1080/02626667.2019.1587562.

Tessier, D. (1990). Behaviour and microstructure of clay minerals. In *Soil colloids and their associations in aggregates* (pp. 387–415). Springer.

Teutschbein, C. & Seibert, J. (2013). Is bias correction of regional climate model (RCM) simulations possible for non-stationary conditions? *Hydrology and Earth System Sciences*, 17(12), 5061–5077.

Teutschbein, C., Grabs, T., Laudon, H., Karlsen, R. H. & Bishop, K. (2018). Simulating streamflow in ungauged basins under a changing climate: The importance of landscape characteristics. *Journal of Hydrology*, 561, 160–178.

Thiboult, A., Anctil, F. & Boucher, M.-A. (2016). Accounting for three sources of uncertainty in ensemble hydrological forecasting. *Hydrology and Earth System Sciences*, 20(5), 1809–1825. Retrieved from: https://doi.org/10.5194/hess-20-1809-2016.

Thober, S., Cuntz, M., Kelbling, M., Kumar, R., Mai, J. & Samaniego, L. (2019). The multiscale routing model mRM v1. 0: simple river routing at resolutions from 1 to 50 km. *Geoscientific Model Development*, 12(6), 2501–2521. Retrieved from: https://doi.org/10.5194/gmd-12-2501-2019.

Tolson, B. A. & Shoemaker, C. A. (2007). Dynamically dimensioned search algorithm for computationally efficient watershed model calibration. *Water Resources Research*, 43(1). Retrieved from: https://doi.org/10.1029/2005WR004723.

TrenberTh, K. (2010). More knowledge, less certainty. *Nature reports climate change*, 4(2), 20–21.

Trenberth, K., Covey, C. & Dai, A. (2018). *Final Report on" Collaborative Research to Narrow Uncertainties in Precipitation and the Hydrological Cycle in Climate Models"*.

Troy, T. J., Wood, E. F. & Sheffield, J. (2008). An efficient calibration method for continental-scale land surface modeling. *Water Resources Research*, 44(9).

Trzaska, S. & Schnarr, E. (2014). A review of downscaling methods for climate change projections. *United States Agency for International Development by Tetra Tech ARD*, 1–42.

Tuller, M. & Or, D. (2003). Hydraulic functions for swelling soils: pore scale considerations. *Journal of hydrology*, 272(1-4), 50–71.

Turcotte, B., Morse, B. & Pelchat, G. (2020). Impact of Climate Change on the Frequency of Dynamic Breakup Events and on the Risk of Ice-Jam Floods in Quebec, Canada. *Water*, 12(10), 2891. Retrieved from: https://doi.org/10.3390/w12102891.

Tyralis, H., Papacharalampous, G. & Langousis, A. (2019). A brief review of random forests for water scientists and practitioners and their recent history in water resources. *Water*, 11(5), 910.

Uhlenbrook, S., Seibert, J., Leibundgut, C. & Rodhe, A. (1999). Prediction uncertainty of conceptual rainfall-runoff models caused by problems in identifying model parameters and structure. *Hydrological Sciences Journal*, 44(5), 779–797.

Van Uytven, E., De Niel, J. & Willems, P. (2020). Uncovering the shortcomings of a weather typing method. *Hydrology and Earth System Sciences*, 24(5), 2671–2686.

Van Vuuren, D. P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hurtt, G. C., Kram, T., Krey, V., Lamarque, J.-F. et al. (2011). The representative concentration pathways: an overview. *Climatic change*, 109(1), 5–31.

Vandewiele, G. & Elias, A. (1995). Monthly water balance of ungauged catchments obtained by geographical regionalization. *Journal of hydrology*, 170(1-4), 277–291.

Vautard, R., Kadygrov, N., Iles, C., Boberg, F., Buonomo, E., Bülow, K., Coppola, E., Corre, L., van Meijgaard, E., Nogherotto, R. et al. (2021). Evaluation of the large EURO-CORDEX regional climate model ensemble. *Journal of Geophysical Research: Atmospheres*, 126(17), e2019JD032344.

Vetter, T., Reinhardt, J., Flörke, M., van Griensven, A., Hattermann, F., Huang, S., Koch, H., Pechlivanidis, I. G., Plötner, S., Seidou, O. et al. (2017). Evaluation of sources of uncertainty in projected hydrological changes under climate change in 12 large-scale river basins. *Climatic Change*, 141(3), 419–433.

Viglione, A. & Blöschl, G. (2009). On the role of storm duration in the mapping of rainfall to flood return periods. *Hydrology and Earth System Sciences*, 13(2), 205–216.

Viglione, A., Chirico, G. B., Woods, R. & Blöschl, G. (2010). Generalised synthesis of space–time variability in flood response: An analytical framework. *Journal of Hydrology*, 394(1-2), 198–212.

Viglione, A., Merz, B., Viet Dung, N., Parajka, J., Nester, T. & Blöschl, G. (2016). Attribution of regional flood changes based on scaling fingerprints. *Water resources research*, 52(7), 5322–5340.

Viney, N. R., Bormann, H., Breuer, L., Bronstert, A., Croke, B. F., Frede, H., Gräff, T., Hubrechts, L., Huisman, J. A., Jakeman, A. J. et al. (2009). Assessing the impact of land use change on hydrology by ensemble modelling (LUCHEM) II: Ensemble combinations and predictions. *Advances in water resources*, 32(2), 147–158. Retrieved from: https://doi.org/10.1016/j.advwatres.2008.05.006.

Volpi, E., Di Lazzaro, M., Bertola, M., Viglione, A. & Fiori, A. (2018). Reservoir effects on flood peak discharge at the catchment scale. *Water Resources Research*, 54(11), 9623–9636.

Von Freyberg, J., Radny, D., Gall, H. E. & Schirmer, M. (2014). Implications of hydrologic connectivity between hillslopes and riparian zones on streamflow composition. *Journal of Contaminant Hydrology*, 169, 62–74.

Vorogushyn, S., Lindenschmidt, K.-E., Kreibich, H., Apel, H. & Merz, B. (2012). Analysis of a detention basin impact on dike failure probabilities and flood risk for a channel-dike-floodplain system along the river Elbe, Germany. *Journal of Hydrology*, 436, 120–131.

Vorogushyn, S., Bates, P. D., de Bruijn, K., Castellarin, A., Kreibich, H., Priest, S., Schröter, K., Bagli, S., Blöschl, G., Domeneghetti, A. et al. (2018). Evolutionary leap in large-scale flood risk assessment needed. *Wiley Interdisciplinary Reviews: Water*, 5(2), e1266.

Wachowicz, L. J., Mote, T. L. & Henderson, G. R. (2019). A rain on snow climatology and temporal analysis for the eastern United States. *Physical Geography*, 1–16.

Wagener, T., McIntyre, N., Lees, M., Wheater, H. & Gupta, H. (2003). Towards reduced uncertainty in conceptual rainfall-runoff modelling: Dynamic identifiability analysis. *Hydrological Processes*, 17(2), 455–476.

Wagener, T. & Wheater, H. S. (2006). Parameter estimation and regionalization for continuous rainfall-runoff models including uncertainty. *Journal of hydrology*, 320(1-2), 132–154.

Wagener, T., Boyle, D. P., Lees, M. J., Wheater, H. S., Gupta, H. V. & Sorooshian, S. (2001). A framework for development and application of hydrological models. *Hydrology and Earth System Sciences*, 5(1), 13–26.

Wagener, T., Wheater, H. & Gupta, H. V. (2004). *Rainfall-runoff modelling in gauged and ungauged catchments*. World Scientific.

Wanders, N. & Wada, Y. (2015). Human and climate impacts on the 21st century hydrological drought. *Journal of Hydrology*, 526, 208–220. Retrieved from: https://doi.org/10.1016/j.jhydrol.2014.10.047.

Wang, C., Graham, R. M., Wang, K., Gerland, S. & Granskog, M. A. (2019). Comparison of ERA5 and ERA-Interim near-surface air temperature, snowfall and precipitation over Arctic sea ice: effects on sea ice thermodynamics and evolution. *The Cryosphere*, 13(6), 1661–1679. Retrieved from: https://doi.org/10.5194/tc-13-1661-2019.

Wang, F., Huang, G., Fan, Y. & Li, Y. (2020). Robust subsampling ANOVA methods for sensitivity analysis of water resource and environmental models. *Water Resources Management*, 34(10), 3199–3217.

Ward, P. J., Jongman, B., Kummu, M., Dettinger, M. D., Sperna Weiland, F. C. & Winsemius, H. C. (2014). Strong influence of El Niño Southern Oscillation on flood risk around the world. *Proceedings of the National Academy of Sciences*, 111(44), 15659–15664.

Warsta, L., Niemi, T. J., Taka, M., Krebs, G., Haahti, K., Koivusalo, H. & Kokkonen, T. (2017). Development and application of an automated subcatchment generator for SWMM using open data. *Urban Water Journal*, 14(9), 954–963. Retrieved from: https://doi.org/10.1080/1573062X.2017.1325496.

Watt, W. E. & Chow, K. A. (1985). A general expression for basin lag time. *Canadian Journal of Civil Engineering*, 12(2), 294–300.

Wechsler, S. (2007). Uncertainties associated with digital elevation models for hydrologic applications: a review. *Hydrology and Earth System Sciences*, 11(4), 1481–1500.

Wei, B.-C., Hu, Y.-Q. & Fung, W.-K. (1998). Generalized leverage and its applications. *Scandinavian Journal of statistics*, 25(1), 25–37.

Wei, Y.-M., Han, R., Liang, Q.-M., Yu, B.-Y., Yao, Y.-F., Xue, M.-M., Zhang, K., Liu, L.-J., Peng, J., Yang, P. et al. (2018). An integrated assessment of INDCs under shared socioeconomic pathways: an implementation of C3IAM. *Natural Hazards*, 92(2), 585–618.

Weigel, A. P., Liniger, M. & Appenzeller, C. (2008). Can multi-model combination really enhance the prediction skill of probabilistic ensemble forecasts? *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography*, 134(630), 241–260.

Weinberg, G. M. (2001). *An introduction to general systems thinking (silver anniversary ed.).* Dorset House Publishing Co., Inc.

Western, A. W., Blöschl, G. & Grayson, R. B. (2001). Toward capturing hydrologically significant connectivity in spatial patterns. *Water Resources Research*, 37(1), 83–97.

Westra, S., Fowler, H., Evans, J., Alexander, L., Berg, P., Johnson, F., Kendon, E., Lenderink, G. & Roberts, N. (2014). Future changes to the intensity and frequency of short-duration extreme rainfall. *Reviews of Geophysics*, 52(3), 522–555.

Westra, S., Alexander, L. V. & Zwiers, F. W. (2013). Global increasing trends in annual maximum daily precipitation. *Journal of climate*, 26(11), 3904–3918.

Wilby, R. L. & Harris, I. (2006). A framework for assessing uncertainties in climate change impacts: Low-flow scenarios for the River Thames, UK. *Water resources research*, 42(2).

Wilks, D. S. (1992). Adapting stochastic weather generation algorithms for climate change studies. *Climatic change*, 22(1), 67–84.

Willems, P., Ntegeka, V., Baguis, P. & Roulin, E. (2010). Climate change impacts on hydrological extremes along rivers and urban drainage systems in Belgium. *Seminarie Belgisch Nationaal Committee voor Geodesie en Geofysica (BNCGG), Date: 2010/04/29-2010/04/29, Location: Leuven*.

Williams, M. W., Erickson, T. A. & Petrzelka, J. L. (2010). Visualizing meltwater flow through snow at the centimetre-to-metre scale using a snow guillotine. *Hydrological processes*, 24(15), 2098–2110.

Wood, A. W., Leung, L. R., Sridhar, V. & Lettenmaier, D. (2004). Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. *Climatic change*, 62(1-3), 189–216.

Wood, E. F., Sivapalan, M., Beven, K. & Band, L. (1988). Effects of spatial variability and scale with implications to hydrologic modeling. *Journal of hydrology*, 102(1-4), 29–47. Retrieved from: https://doi.org/10.1016/0022-1694(88)90090-X.

Woods, R. & Sivapalan, M. (1999). A synthesis of space-time variability in storm response: Rainfall, runoff generation, and routing. *Water Resources Research*, 35(8), 2469–2485.

Wright, D. P., Thyer, M. & Westra, S. (2015). Influential point detection diagnostics in the context of hydrological model calibration. *Journal of Hydrology*, 527, 1161–1172.

Wright, D. P., Thyer, M., Westra, S. & McInerney, D. (2018). A hybrid framework for quantifying the influence of data in hydrological model calibration. *Journal of hydrology*, 561, 211–222.

Wright, D. P., Thyer, M., Westra, S., Renard, B. & McInerney, D. (2019). A generalised approach for identifying influential data in hydrological modelling. *Environmental modelling & software*, 111, 231–247.

Yang, W., Andréasson, J., Phil Graham, L., Olsson, J., Rosberg, J. & Wetterhall, F. (2010). Distribution-based scaling to improve usability of regional climate model projections for hydrological climate change impacts studies. *Hydrology Research*, 41(3-4), 211–229.

Yang, W., Long, D. & Bai, P. (2019a). Impacts of future land cover and climate changes on runoff in the mostly afforested river basin in North China. *Journal of Hydrology*, 570, 201–219. Retrieved from: https://doi.org/10.1016/j.jhydrol.2018.12.055.

Yang, X., Magnusson, J. & Xu, C.-Y. (2019b). Transferability of regionalization methods under changing climate. *Journal of Hydrology*, 568, 67–81.

Yin, J., Gentine, P., Zhou, S., Sullivan, S. C., Wang, R., Zhang, Y. & Guo, S. (2018). Large increase in global storm runoff extremes driven by climate and anthropogenic changes. *Nature communications*, 9(1), 1–10. Retrieved from: https://doi.org/10.1038/s41467-018-06765-2.

Zehe, E., Elsenbeer, H., Lindenmaier, F., Schulz, K. & Blöschl, G. (2007). Patterns of predictability in hydrological threshold systems. *Water Resources Research*, 43(7).

Zehe, E. & Blöschl, G. (2004). Predictability of hydrologic response at the plot and catchment scales: Role of initial conditions. *Water Resources Research*, 40(10).

Zehe, E., Blume, T. & Blöschl, G. (2010). The principle of 'maximum energy dissipation': a novel thermodynamic perspective on rapid water flow in connected soil structures. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1545), 1377–1386.

Zehe, E., Ehret, U., Pfister, L., Blume, T., Schröder, B., Westhoff, M., Jackisch, C., Schymanski, S. J., Weiler, M., Schulz, K. et al. (2014). HESS Opinions: From response units to functional units: a thermodynamic reinterpretation of the HRU concept to link spatial organization and functioning of intermediate scale catchments. *Hydrology and Earth System Sciences*, 18(11), 4635–4655.

Zhang, L., Nan, Z., Yu, W., Zhao, Y. & Xu, Y. (2018). Comparison of baseline period choices for separating climate and land use/land cover change impacts on watershed hydrology using distributed hydrological models. *Science of the Total Environment*, 622, 1016–1028.

Zhang, W. & Montgomery, D. R. (1994). Digital elevation model grid size, landscape representation, and hydrologic simulations. *Water resources research*, 30(4), 1019–1028.

Zhang, X., Srinivasan, R., Zhao, K. & Liew, M. V. (2009). Evaluation of global optimization algorithms for parameter calibration of a computationally intensive hydrologic model. *Hydrological Processes: An International Journal*, 23(3), 430–441.

Zhang, Y., Viglione, A. & Blöschl, G. (2022). Temporal scaling of streamflow elasticity to precipitation: a global analysis. *Water Resources Research*, 58(1), e2021WR030601.

Zhao, F., Wu, Y., Qiu, L., Sun, Y., Sun, L., Li, Q., Niu, J. & Wang, G. (2018). Parameter uncertainty analysis of the SWAT model in a mountain-loess transitional watershed on the Chinese Loess Plateau. *Water*, 10(6), 690. Retrieved from: https://doi.org/10.3390/w10060690.

Zheng, F., Westra, S., Leonard, M. & Sisson, S. A. (2014). Modeling dependence between extreme rainfall and storm surge to estimate coastal flooding risk. *Water Resources Research*, 50(3), 2050–2071.

Zhong, R., He, Y. & Chen, X. (2018). Responses of the hydrological regime to variations in meteorological factors under climate change of the Tibetan plateau. *Atmospheric Research*, 214, 296–310.

Zhou, Z., Smith, J. A., Yang, L., Baeck, M. L., Chaney, M., Ten Veldhuis, M.-C., Deng, H. & Liu, S. (2017). The complexities of urban flood response: Flood frequency analyses for the Charlotte metropolitan region. *Water Resources Research*, 53(8), 7401–7425. Retrieved from: https://doi.org/10.1002/2016WR019997.

Zhuan, M.-J., Chen, J., Shen, M.-X., Xu, C.-Y., Chen, H. & Xiong, L.-H. (2018). Timing of human-induced climate change emergence from internal climate variability for hydrological impact studies. *Hydrology Research*, 49(2), 421–437.

Ziegler, A. & König, I. R. (2014). Mining data with random forests: current options for real-world applications. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 4(1), 55–63.

Zscheischler, J. & Seneviratne, S. I. (2017). Dependence of drivers affects risks associated with compound events. *Science Advances*, 3(6), e1700263.

Zscheischler, J., Westra, S., Van Den Hurk, B. J., Seneviratne, S. I., Ward, P. J., Pitman, A., AghaKouchak, A., Bresch, D. N., Leonard, M., Wahl, T. et al. (2018). Future climate risk from compound events. *Nature Climate Change*, 8(6), 469.