RESEARCH ARTICLE

Correspondence to:
P. Lucas-Picher,
p.lucas.picher@gmail.com

Application of a High-Resolution Distributed Hydrological Model on a U.S.-Canada Transboundary Basin: Simulation of the Multiyear Mean Annual Hydrograph and 2011 Flood of the Richelieu River Basin

Philippe Lucas-Picher[1,2] ID, **Richard Arsenault[1]** ID, **Annie Poulin[1]** ID, **Simon Ricard[3,4]** ID, **Simon Lachance-Cloutier[3]** ID, **and Richard Turcotte[3,5]** ID

[1]Department of Construction Engineering, École de Technologie Supérieure, Montreal, Quebec, Canada, [2]Now at Groupe de Météorologie de Grande Échelle et Climat, Centre National de Recherches Météorologiques, Météo-France, Toulouse, France, [3]Ministère de l'Environnement et de la Lutte contre les Changements Climatiques du Québec (MELCC), Québec, Quebec, Canada, [4]Department of Civil and Water Engineering, Université Laval, Québec, Quebec, Canada, [5]Ouranos/Consortium on Regional Climatology and Adaptation to Climate Change, Montréal, Quebec, Canada

Abstract During spring 2011, an extreme flood occurred along the Richelieu River located in southern Quebec, Canada. The Richelieu River is the last section of the complex Richelieu basin, which is composed of the large Lake Champlain located in a valley between two large mountains. Previous attempts in reproducing the Richelieu River flow relied on the use of simplified lumped models and showed mixed results. In order to prepare a tool to assess accurately the change of flood recurrences in the future, a state-of-the-art distributed hydrological model was applied over the Richelieu basin. The model setup comprises several novel methods and data sets such as a very high resolution river network, a modern calibration technique considering the net basin supply of Lake Champlain, a new optimization algorithm, and the use of an up-to-date meteorological data set to force the model. The results show that the hydrological model is able to satisfactorily reproduce the multiyear mean annual hydrograph and the 2011 flow time series when compared with the observed river flow and an estimation of the Lake Champlain net basin supply. Many factors, such as the quality of the meteorological forcing data, that are affected by the low density of the station network, the steep terrain, and the lake storage effect challenged the simulation of the river flow. Overall, the satisfactory validation of the hydrological model allows to move to the next step, which consists in assessing the impacts of climate change on the recurrence of Richelieu River floods.

Plain Language Summary In order to study the 2011 Richelieu flood and prepare a tool capable of estimating the effects of climate change on the recurrence of floods, a hydrological model is applied over the Richelieu basin. The application of a distributed hydrological model is useful to simulate the flow of all the tributaries of the Richelieu basin. This new model setup stands out from past models due to its distribution in several hydrological units, its high-resolution river network, the calibration technique, and the high-resolution weather forcing data set used to drive the model. The model successfully reproduced the 2011 Richelieu River flood and the annual hydrograph. The simulation of the Richelieu flow was challenging due to the contrasted elevation of the Richelieu basin and the presence of the large Lake Champlain that acts as a reservoir and attenuates short-term fluctuations. Overall, the application was deemed satisfactory, and the tool is ready to assess the impacts of climate change on the recurrence of Richelieu River floods.

1. Introduction

During spring 2011, an unprecedented flood hit the Richelieu River (Figure 1), located in the southern province of Quebec, Canada, and Lake Champlain, located in Vermont and New York states in northeastern United States (Figure 2). A few studies analyzed this flood thoroughly and revealed that many factors contributed to its magnitude (Lucas-Picher et al., 2015; Riboust & Brissette, 2016; Saad et al., 2016); namely, the 2011 flood was caused by a combination of an important snowpack during winter and extreme
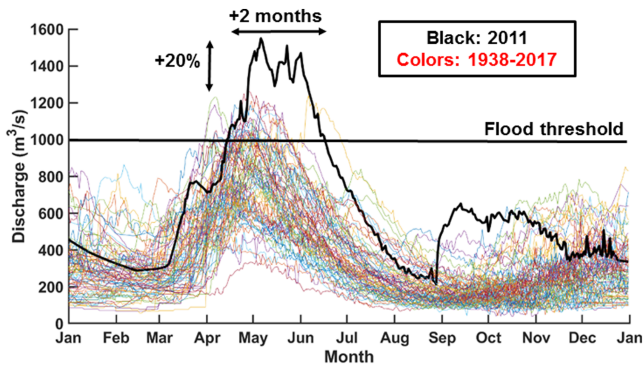
Figure 1. Daily flow of the Richelieu River at the Fryers hydrometric station for the period 1938–2017. Each year is presented with different colors while 2011 is in black.

precipitation in spring, and late spring sustained warm conditions maintaining snowmelt during the freshet. These main contributing flood factors led to an increase of inflows into Lake Champlain, resulting in a rapid increase of water levels. Consequently, the lake outflow toward the Richelieu River increased as well. On 14 April 2011, the flow of the Richelieu River reached the flood threshold of 1,000 m^3/s at the Fryers hydrometric station (see Figures 2 and 5 for the location of the station) located close to the town of St-Jean-Sur-Richelieu, Quebec. On 6 May 2011, the Richelieu River reached a record-breaking flow of 1,550 m^3/s. In addition, the 2011 peak flow was also affected by a prolonged wind surge that increased the water level in the lake north (i.e., the outlet), increasing even more the Richelieu River flow. Finally, the flow remained above 1,000 m^3/s until 16 June 2011. It is also interesting to note the sharp rise of the Richelieu River discharge at the end of August 2011 (Figure 1) that remained anomalously high in September and October. This behavior was linked to the large precipitation that occurred during the passage of the tropical storm Irene over the Richelieu basin.

Overall, the 2011 spring flood largely exceeded all previous floods in terms of intensity, with a peak river flow of 20% above the previous record, and duration, lasting for more than 2 months above the flood threshold (Figure 1). Hundreds of homes were flooded for more than 2 months, requiring waves of evacuation and military involvement to support the effort (Riboust & Brissette, 2016). The flood was ranked third among Canada's top 10 weather stories of 2011 and was one of Quebec's longest-lasting natural disasters ever, wearing out residents, physically and mentally (Lucas-Picher et al., 2015). Currently, efforts are ongoing to better understand the causes of the flood and to estimate how climate change will affect the probabilities of such a flood of reoccurring in the future (International Joint Commission [IJC], 2017).

The Richelieu basin (Figure 2) is a large (23,881 km^2) and complex transboundary basin mainly located over northeastern United States. From the outlet of Lake Champlain, the water flows northward through the Richelieu River toward the Saint Lawrence River located in the southern portion of Quebec, Canada. More precisely, the Richelieu basin originates in small streams in the Adirondack Mountains (State of New York) and Green Mountains (State of Vermont) flowing downward to Lake Champlain, which is located in a long (200 km) and narrow (25 km) valley. Overall, 90% of the water of the Richelieu basin arises from the Lake Champlain basin (21,325 km^2), and the flow in the Richelieu River is mainly controlled by the water level of the lake. For more details about the Richelieu River and Lake Champlain, the reader is referred to Shanley and Denner (1999).

With high mountains (more than 50 peaks above 1,200 m) bordering a deep valley (close to sea level) where Lake Champlain is located, different weather conditions take place over the Richelieu basin (Figure 3). The weather in the mountains is cold and snowy in winter, while it is warmer and dryer in the valley. When the warm and moist air from the Atlantic Ocean and southern United States is lifted due to the high peaks of the Adirondack and Green Mountains, clouds and orographic precipitation (rain/snow) are generated. With these different weather conditions occurring over only a few kilometers, the Richelieu basin experiences many regional and local weather conditions. To capture adequately historical conditions, there is a need to have either: climate models with fine horizontal and vertical resolutions or a dense weather station network to facilitate the production of gridded data set (Lucas-Picher et al., 2015). Due to its location in midlatitudes, winters in the Richelieu basin are cold with mean minimum temperature below −10 °C, while summers are warm with mean maximum temperature
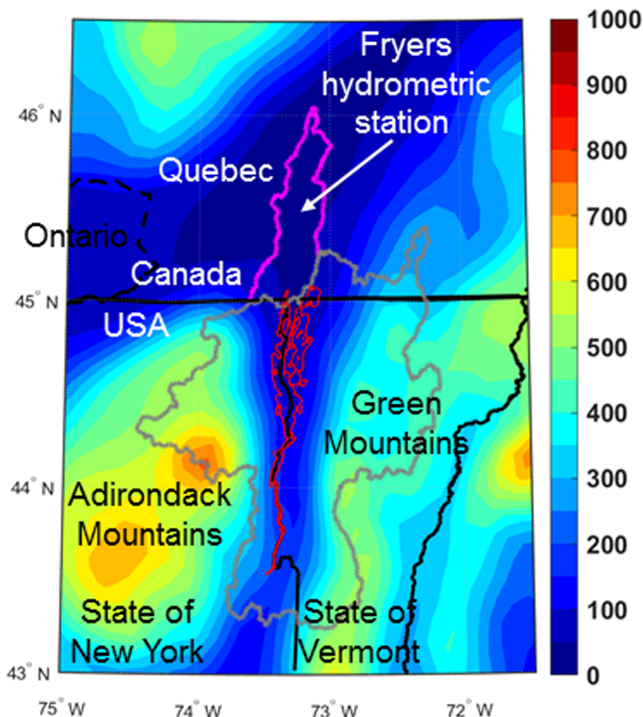

Figure 2. Topography (m) and contour of the Richelieu basin (magenta and gray), Lake Champlain basin (gray), and Lake Champlain (red).
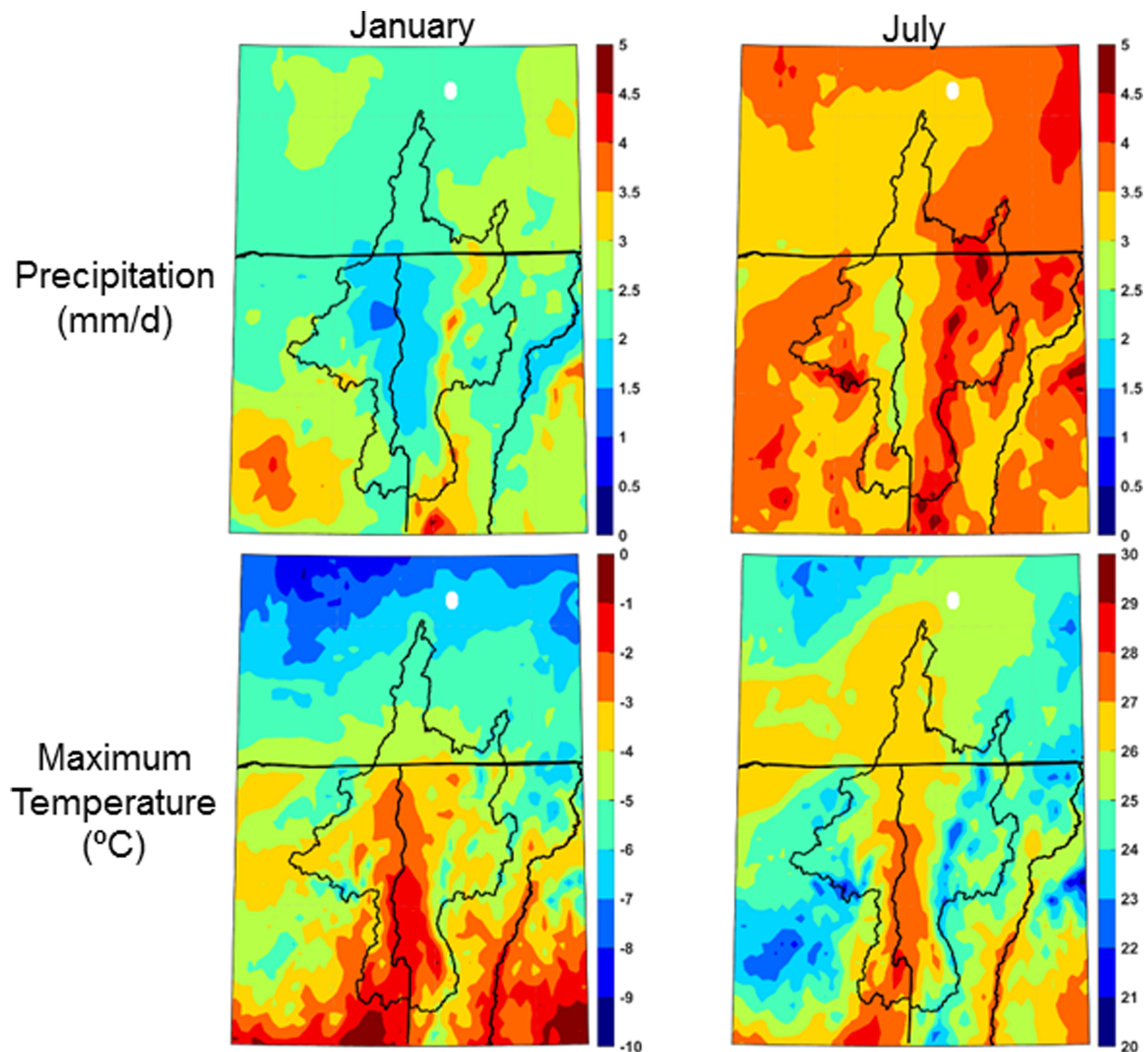
Figure 3. 1981–2010 mean precipitation (top row) and mean maximum temperature (bottom row) in January (left column) and July (right column) over the Richelieu basin and its surroundings from the ~7-km Livneh et al. (2015) meteorological data set.

above 20 °C (Figure 4). Finally, precipitation is a bit higher in summer, with a mean around 3.5 mm/day, compared to winter, with a mean around 2.5 mm/day (Figure 4).

In response to the above-mentioned climate conditions and to the large surface area of the river basin, the annual hydrograph displays an important spring freshet (Figure 1). Late summer and winter are affected by low flows due to high evapotranspiration in summer and snow accumulation in winter reducing runoff and inflows to the rivers and the lake. With colder and rainy conditions in fall, the river flow usually rises after the summer low flows. Finally, the flows of the Lake Champlain tributaries are regulated at few locations, but the effects are localized and overall; the flows of the main rivers can be considered natural.

The simulation of the Richelieu basin with hydrological models, and especially the 2011 flood, involves many different challenges. With Lake Champlain's large size (1,269 km^2, 170 km long by 20 km wide), providing most of the water to the Richelieu River, river flow modeling of the basin becomes a complex task. The lake itself acts as a large reservoir that attenuates short-term fluctuations from upstream rivers through lake routing (Riboust & Brissette, 2015). With a mean depth of 20 m, the lake contains around 26 km^3 of water that flows at a mean rate of 375 m^3/s through the Richelieu River. The difference of temperature between winter and summer can reach 50 °C, and during long periods in spring and fall, the temperature often oscillates above and below 0 °C on a daily basis. In addition, with the valley located between
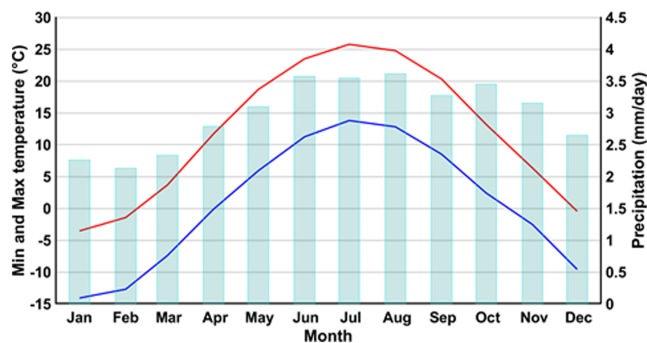
Figure 4. 1981–2010 mean annual cycle for precipitation (cyan bars, right axis), minimum temperature (blue line, left axis), and maximum temperature (red line, left axis) over the Richelieu basin from the ~7-km Livneh et al. (2015) meteorological data set.

two high mountain ranges, many different regional and local climate conditions that take place have to be measured and collected to produce an accurate meteorological forcing of the hydrological simulation. However, the low population density and the steep elevations of the region explain why only a few weather stations can be found in the region, making it difficult to find good quality high-resolution representations of the weather conditions. Errors in the daily precipitation and temperature forcing data sets can affect the calibration and the validation of the hydrological model in a region with a low density of weather stations (Essou et al., 2017) such as the Richelieu basin. With long and cold winters, an important snowpack builds up over the basin (usually close to 1 m basin-averaged in March), mainly over the mountains, and melts in spring. Moreover, due to the large heat capacity of the water from the lake, the surroundings of the lake are milder in winter and cooler in summer than the regions away from the lake.

The flow of the Richelieu River has been simulated using hydrological models in a few studies. Boyer et al. (2010) used a lumped hydrological model to study the impact of climate change on the tributaries of the Saint Lawrence River, with the Richelieu River being one of them. Once the hydrological model calibrated, Boyer et al. (2010) succeeded in reproducing the mean spring high flow of the Richelieu River during the reference period but overestimated the summer mean low flow by more than 50%. A few years later, Riboust and Brissette (2015) used the same lumped model but constructed a reservoir model for the lake in a semidistributed fashion and simulated the Richelieu River flow following a relationship between the lake level and the river flow. They obtained a Nash-Sutcliffe Efficiency criterion of 0.82 between simulated and observed Richelieu River flows with a slight tendency to underestimate high flows and to overestimate low flows. In a study highlighting the added value of regional climate model (RCM) simulations driven at their boundaries by a reanalysis, Lucas-Picher et al. (2015) forced the hydrological model setup of Riboust and Brissette (2015) with the outputs from two RCMs. Despite clear added value of the RCM simulations, the hydrological model simulations driven by the RCMs could not reproduce the 2011 peak flow, likely due to the underestimation of precipitation in spring. As in the previous studies, the multiyear mean annual hydrographs in Lucas-Picher et al. (2015) also suffered from underestimation of mean spring flows and overestimation of mean low flows. Finally, the flow of the rivers of the Richelieu basin is also simulated in the Hydroclimatic Atlas of Southern Quebec (Direction de l'expertise hydrique [DEH], 2018a) but without specific meteorological data in the United States, and the resolution of the hydrological network is rather coarse. These shortcomings affect the quality of the simulated flow, explaining why the streams of the Richelieu basin are masked in the final version of the Atlas available on the web. In a comparison of the river flow simulation using three hydrological models of different levels of complexity, Ludwig et al. (2009) and Velazquez et al. (2013) emphasized the positive impact of using a distributed hydrological model in a contrasted watershed.

Taking advantage of recent advances in hydrological modeling, faster computers, and new and improved gridded observations, the current paper addresses some limitations of lumped models by applying a high-resolution distributed hydrological model over the Richelieu basin. Many advanced and novel techniques were used to improve the simulation of the flow of the tributaries of Lake Champlain and Richelieu River. First, the use of a high-resolution distributed model should help in representing the different elevations and features of the basin and improve the simulation of river flows throughout the hydrographic network. Then, the hydrological model was calibrated using a recent two-step technique (local followed by global calibration; Ricard et al., 2013) that allows spatial homogeneity of the parameters. During the calibration, the identification of parameters was performed with a novel hybrid optimization approach (Huot et al., 2019) adapted for computationally intensive models, substantially reducing the number of model evaluations required to find the optimum parameters. In addition, the hydrological simulations benefited from a new ~7-km observationally based gridded data set where precipitation was adjusted to improve orographic effects (Livneh et al., 2015). This up-to-date hydrological model setup over the Richelieu basin was evaluated through its ability to reproduce the multiyear mean annual hydrograph and the 2011 flood. Finally, the current application of the hydrological model constitutes the basis of a study that will assess whether climate
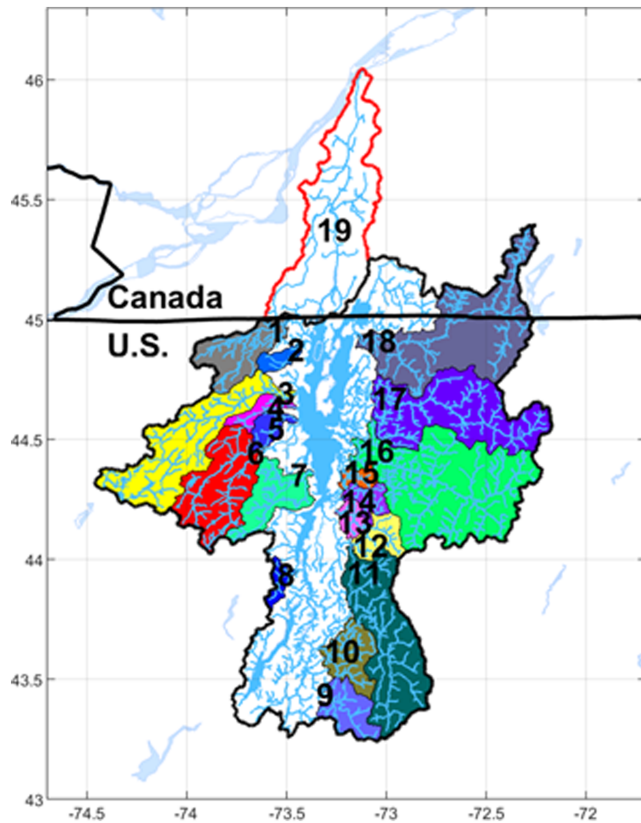
Figure 5. Hydrological network used by HYDROTEL. The 18 subbasins of Lake Champlain used for the calibration are indicated with different colors and numbers. The location of Fryers station is indicated with #19.

change will affect Richelieu River floods. This study is part of a larger International Joint Commission project that will provide recommendations on flood mitigation measures that should remain relevant under projected climate change (IJC, 2017).

In the next section, the full methodology, including the hydrological model and the model setup over the region of interest, is described. Then, section 3 presents the results of the calibration and the validation of the hydrological setup through an evaluation of a long hydrological simulation with observed river flows and lake net basin supply (NBS). Section 4 introduces a discussion of the results, and finally, section 5 presents the conclusions.

2. Methodology and Experimental Setup

2.1. Hydrological Model HYDROTEL and Model Setup Over the Richelieu Basin

Since the early 1980s, the hydrological model HYDROTEL (Fortin et al., 1995, 2001; Turcotte et al., 2007) has been developed by a group of researchers at the "Institut National de Recherche Scientifique" (INRS) in Quebec City, Canada. It has been applied on several watersheds—especially over the Quebec province—to study, among other things, the impacts of hydrological model structure in climate change impact assessment studies (Chen et al., 2011; Poulin et al., 2011; Velazquez et al., 2013), the simulation of the effects of wetlands on watershed hydrology under current and future climate and land cover conditions (e.g., Blanchette et al., 2019; Fossey et al., 2015, 2016; Fossey & Rousseau, 2016a, 2016b), the simulation of snow water equivalent (SWE; Oreiller et al., 2014), and the assessment of future trends in low flows (Foulon et al., 2018). Since many years, HYDROTEL is the main hydrological model used operationally at the *Direction de l'Expertise Hydrique* (Québec's Water Expertise Direction) for short-term flow forecasting (Turcotte et al., 2004) and for simulating the impacts of climate change on river flows as part of the Hydroclimatic Atlas of Southern Quebec (DEH, 2018a). Recently, the model was implemented for short-term operational hydrological forecasting using ensemble Kalman filtering for the Yukon Energy Corporation in northwestern Canada (Samuel et al., 2019). It is noteworthy that the model had previously been used to assess the added value of using ensemble Kalman filtering to improve operational streamflow forecasts during the snow accumulation and melt periods (Abaza et al., 2015) and during the rainfall-runoff production periods (Abaza et al., 2014).

HYDROTEL is a spatially distributed hydrological model and is predominantly (although not fully) physically based. A given watershed is divided in several simulation units called relatively homogeneous hydrological units (RHHUs) and river reaches. The user, based on the desired level of complexity, controls the numbers of RHHU subdivisions and river reaches according to the specified hydrological network discretization. River reaches are one-dimensional elements that support streamflow simulation processes, whereas RHHUs are basic elements for all rainfall-runoff simulation processes. Often referred to hydrologic response units (HRU) in the literature, RHHUs represent either elementary subwatersheds or hillslopes, depending on the user's interest (Noël et al., 2014), taking into account the spatial variability of topography, land cover, soil types, and meteorological variables (Lavigne et al., 2004). In this study, since only one basin is modeled, the network discretization was pushed to a high resolution where the RHHUs and reaches were determined up to the Strahler order number 2. According to this configuration, the entire Richelieu basin is composed of 8,032 RHHUs (i.e., elementary subwatersheds) and 3,208 river segments (Figure 5). The RHHU mean surface area is 3.1 km^2.

HYDROTEL has submodel options for the simulation of snow accumulation and melt, potential evapotranspiration (PET), soil moisture content of each RHHU, runoff generation, and horizontal flow (river routing). Simulations can also be run at daily or subdaily time steps. In this study, we selected the

mixed degree-day snow submodel following an energy budget approach (Turcotte et al., 2007), while PET is based on the Linacre equation (Linacre, 1977). The Linacre equation was selected following a series of tests, which determined that it outperformed the other available PET formulations. The moisture content in the soil column is modeled by applying the continuity equation to three layers from which surface run-off, interflow, and baseflow are generated, while the vertical fluxes are calculated using the Buckingham-Darcy equation. The river flow is modeled with the kinematic wave equation, and the temporal discretization (time step) is 24 hr. The HYDROTEL model needs to be driven with daily minimum and maximum temperatures and precipitation. For more details on the model, the reader is referred to Fortin et al. (2001) and Turcotte et al. (2007).

In order to perform a hydrological simulation with HYDROTEL, different physiographic fields such as watershed boundary, land cover, soil type, river network, RHHU delineation, flow direction, and elevation are needed. Those fields are determined from PHYSITEL (Turcotte et al., 2001), which is a specialized geographic information system that supports the application of HYDROTEL over a specific basin. The elevation of the Richelieu basin used by HYDROTEL is shown in Figure 6a. The elevation of the RHHU was calculated based on a 100-m aggregation of a digital elevation model at 10-m resolution over Quebec (BDTQ - Ministère des Ressources naturelles et de la Faune, 2008) and 30-m resolution over the United States (NED – United Geological of the United States, 1999). From that figure, Lake Champlain, bordered by the Adirondack and Green Mountains, can be clearly seen. The soil type was estimated using the 30″ resolution database of Shangguan et al. (2014), while the land cover was estimated from different sources over the Quebec province and the 300-m resolution Globcover2009 (Bontemps et al., 2011) in the United States. The soil type (Figure 6b) is mainly loam over the United States and clay over Canada. The land cover (Figure 6c) is mainly deciduous forest in the United States and cropland in Canada. Finally, due to the high resolution, most RHHUs are smaller than 25 km^2 where the elevation is high in the United States and close to Lake Champlain; meanwhile in Canada, the elevation is low, and RHHUs are often larger than 25 km^2 (Figure 6d).

HYDROTEL treats lakes and river streams differently. The water volume in lakes is simulated following a simple reservoir model based on the continuity equation:

$$A \, \Delta h = (Q_1 - Q_0) \, \Delta t, \tag{1}$$

where A is the lake area, h is the change in water level, Q_1 is the lake inflow, t is the time step, and Q_0 is the lake outflow. The flow at the lake's outlet is computed following a level flow curve:

$$Q_0 = C \, h^k, \tag{2}$$

where C and k are tunable parameters and h corresponds to the level of the lake, which is a state variable. By default, HYDROTEL uses predefined values for C and k. These coefficients can however be adjusted when river flow data are available downstream of the lake section. In the present setup, those two parameters have a major impact on the simulated flow of the Richelieu River by controlling the lake storage and routing effects. In order to improve the simulated flow of the Richelieu River, the default values of the coefficients C and k were replaced by values calibrated using the observed and simulated flow at the Fryers hydrometric station. Different values of C and k were successively tested. The values $C = 250$ and $k = 1.4$ leading to the smallest RMSE between observed and simulated Richelieu flow over the period 1979–2009 were retained (DEH, 2018b). Moreover, HYDROTEL is a hydrological model that can be used to simulate river flow but not the water level and the extent of inundation. Such process is simulated by a hydraulic model using the simulated river flow as forcing. This will be done at a later stage of the IJC project (IJC, 2017) to evaluate the damages and the economic costs of floods and to evaluate potential adaptation solutions to mitigate flood risks.

2.2. Observationally Based Gridded Observations and River Flow Measurements

HYDROTEL was forced with the spatially comprehensive meteorological data set of Livneh et al. (2015) covering Southern Canada, the conterminous United States, and Mexico. This data set comprises observed daily maximum and minimum temperatures as well as precipitation data interpolated on a 1/16° (~7 km) grid originally for the period 1950–2013 but now extended for the period 1915–2015. Because of the consistent
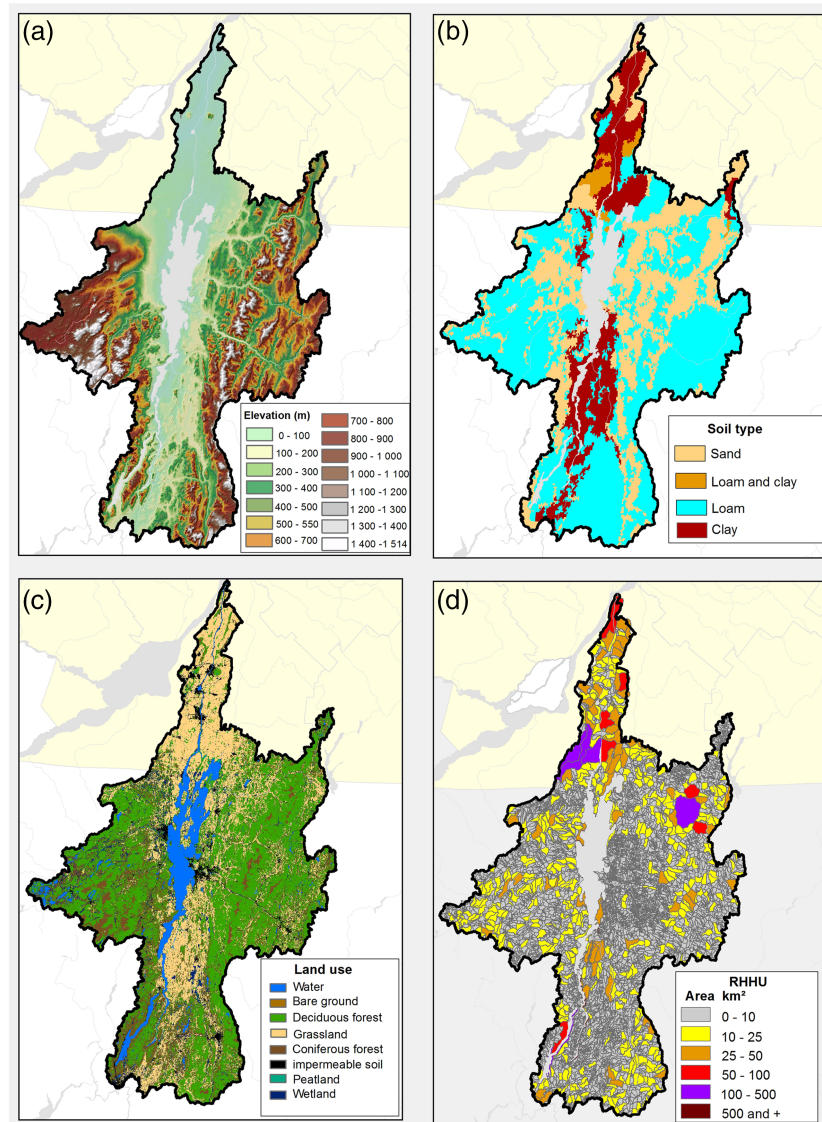
Figure 6. Physiographic fields of the hydrological model HYDROTEL over the Richelieu basin: (a) elevation, (b) soil type, (c) land cover, and (d) RHHU spatial distribution and size.

gridding methodology, the data set of Livneh et al. (2015) reduces transboundary discontinuities as compared to commonly used reanalysis products, making it suitable for estimating large-scale hydrometeorological phenomena.

Daily river flow measurements gathered by the U.S. Geological Survey (USGS) were used to calibrate the hydrological model. Since approximately 90% of the water at the Richelieu River outlet originates from Lake Champlain and to avoid lake storage and routing effects downstream of the lake, a water budget approach was selected to calibrate the model, using 18 gauged stations of different subwatersheds of the lake. The 18 gauged stations are the same as those used by Riboust and Brissette (2015) and are described in Shanley and Denner (1999). The list of stations is indicated in Table 2, and their spatial distribution is shown in Figure 5. With a common coverage of the period 1992–2013 for all the stations, the calibration was performed for the 1992–2003 period, while the validation was done using the 2004–2013 period. The cumulative drainage area of the 18 stations corresponds to 73% of the total water inflow of Lake Champlain.

Table 1
List of HYDROTEL Calibrated Parameters

Parameter	Description	Unit
1	Depth of the first two soil layers	m
2	Depth of the third soil layer	m
3	Coefficient of recession of the third soil layer	m/hr
4	Melt temperature threshold of the snowpack	°C
5	Melt rate of the snowpack	mm/°C/day
6	Potential evapotranspiration multiplication factor	[]
7	Rain to snow temperature threshold	°C

Note. Three parameters are linked to the flow of water in the soil, two parameters to the snow module, one parameter is linked to the potential evapotranspiration, and one parameter is linked to the determination of the precipitation phase.

2.3. Calibration of the Hydrological Model

The calibration of HYDROTEL has been the subject of three studies, namely, those of Turcotte et al. (2003) establishing the basis of a structured manual calibration, Bouda et al. (2014) comparing the performance of this strategy with that of an automatic calibration, and Bouda et al. (2012) conducting an uncertainty analysis of the model. In the current setup, HYDROTEL has seven calibration parameters listed in Table 1. While most of the land cover of the basin is classified as deciduous forest, other parts of the basin are classified as open areas or coniferous forest. One argument in favor of not using land cover dependencies for the snow model is, as our results suggest, that local calibration already show signs of over-parametrization. Thus, increasing the complexity with new parameters is unlikely to improve significantly the flow simulation over the validation period. Given the proportion of deciduous forest, it would also be difficult to calibrate the parameters specifically for this land cover.

For the global calibration, it might have been of interest to use distinct parameters for open areas and forested area; as contrary to the local calibration, the global calibration does not show signs of over-parametrization. On the other hand, we see from Figure 9 that the model already performs quite well during the freshet period. Thus, only little improvements would have been possible using slightly different snowmelt parameters that will mostly affect the timing of the melt between different land covers. The different steps of the calibration are described later on. The main steps of the entire methodology are shown schematically in Figure 7.

The calibration was performed in two steps. The first step consisted of a "local calibration" to find the seven optimal parameter values for each subwatershed. The selected performance indicator was the Kling-Gupta efficiency (KGE) criterion (Gupta et al., 2009) computed using the daily simulated and observed of river flows. In 2012, Kling et al. (2012) proposed an updated KGE that allows a better interpretation of the bias, variance, and correlation terms. It would have been preferable to use the 2012 KGE version, but this modification is unlikely to change the overall results. The novel hybrid optimization algorithm DDS-MADS (Huot et al., 2019), combining the dynamically dimensioned search (DDS; Tolson & Shoemaker, 2007) and the mesh adaptive direct search (MADS; Audet & Dennis, 2006), was designed to perform an efficient optimization by reducing the number of model evaluations while ensuring quality and robustness of the final parameter sets, including proof of convergence in at least a local minimum. DDS-MADS is particularly useful for this model setup that is highly computationally demanding because of the fine resolution of the simulated watershed that includes numerous river segments and RHHUs. The calibration was performed for the 1992–2003 period where the first year was considered as a spin-up period and was removed from the computation of the KGE to allow a good initialization and a well-balanced model.

The second step consisted of a "global calibration" using the performance indicators of the local calibration as targets to find one set of seven optimal parameter values for the entire Lake Champlain basin. This two-step calibration method, using a local then a global calibration, has been validated and described by Ricard et al. (2013), Haghnegahdar et al. (2014), and Gaborit et al. (2015). A global calibration is useful to identify a single set of seven parameter values for a large region, to ensure the spatial consistency of the distributed parameters where all the basins of the same region are modeled uniformly and to optimize the hydrological simulation of ungauged basins. In this study, the objective function (OF) to be minimized by the optimization algorithm was

$$\text{OF} = \sum_{i=1 \text{ to } 18} \text{area}_i * ((1\text{-KGE\_Global}_i)/(1\text{-KGE\_Local}_i)), \qquad (3)$$

where KGE_Global_i corresponds to the KGE computed using the global parameters and KGE_Local_i corresponds to the KGE computed using the local parameters for the basin i from 1 to 18. The value $(1\text{-KGE\_Global}_i)/(1\text{-KGE\_Local}_i)$ of each basin is multiplied by the surface area of the basin to give larger
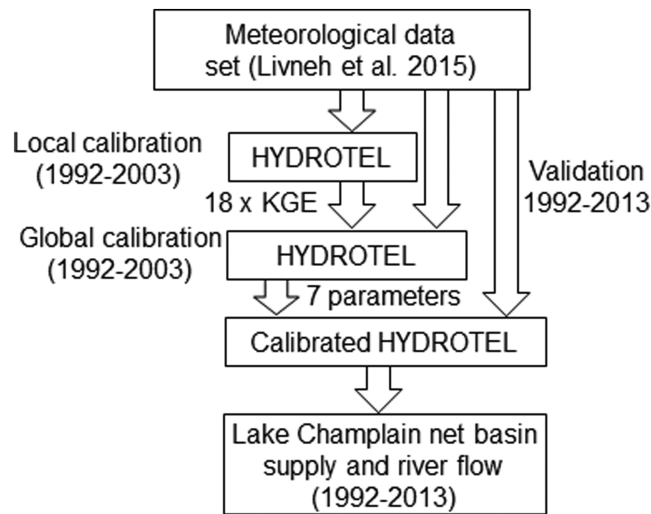
Figure 7. Flow chart showing the main steps of the methodology.

weights to the larger basins. A global calibration was done for the same period as for the local calibration (i.e., 1992–2003), also using the first year as spin-up.

3. Results

3.1. Comparison of the Performance of the Local and Global Calibrations

Table 2 compares the performance indicators (KGE) using the parameters obtained through the local and global parameters over the calibration (1992–2003) and validation period (2004–2013). Over the calibration period (1992–2003), the KGEs using the parameters of the local calibration are on average greater than that of the global calibration with respective average values of 0.83 and 0.75. This is an expected result as in the local calibration, each basin is treated separately, which leads to 18×7 parameters being calibrated, while the global optimization requires only seven parameters. Thus, assuming the convergence of the optimization algorithm, the KGEs using the parameters of the local calibrations should be greater or equal to the KGEs using the parameters of the global calibration. This is the case for 17 of the 18 basins. However, for the remaining basin, the KGE using the parameters of the global calibration outperforms the KGE of the parameters of the local calibration indicating in all likelihood that the optimization algorithm did not completely converge toward the optimal solution. Since the difference of KGE is small, the calibration for the remaining basin was not repeated. Both the local and global calibrations tend to perform slightly better over large basins. This can be partly explained by the reaction time of such basins, which is probably more coherent with the daily simulation time step of the model.

An interesting result arises from comparing performances of the calibration period against the validation period in Table 2. It can be seen that local parameters are related to a substantial loss in performance from the calibration (KGE = 0.83) to the validation (KGE = 0.73) period. On the other hand, the loss of performance using the global parameters is less significant with an average KGE value of 0.75 for the calibration

Table 2
Performance of HYDROTEL Using the KGE Over the Calibration and Validation Periods for the 18 Gauged Rivers Flowing Toward Lake Champlain

#	USGS site number	River basin	Drainage area (km^2)	Local		Global	
				Calibration (1992–2003)	Validation (2004–2013)	Calibration (1992–2003)	Validation (2004–2013)
1	"4271500"	Great Chazy	629.37	0.86	0.72	0.68	0.82
2	"4271815"	Little Chazy	129.5	0.82	0.72	0.5	0.6
3	"4273500"	Saranac	1,574.71	0.89	0.8	0.85	0.82
4	"4273700"	Salmon River	163.17	0.76	0.72	0.64	0.68
5	"4273800"	Little Ausable	173.53	0.76	0.66	0.61	0.7
6	"4275500"	Ausable	1,155.13	0.85	0.75	0.84	0.77
7	"4276500"	Boquet	699.3	0.84	0.74	0.80	0.71
8	"4276842"	Putnam Creek	132.09	0.87	0.62	0.71	0.77
9	"4280450"	Mettawee	432.53	0.85	0.59	0.72	0.33
10	"4280000"	Poultney	484.33	0.87	0.63	0.79	0.74
11	"4282500"	Otter Creek	1,626.51	0.85	0.68	0.81	0.68
12	"4282525"	New Haven	297.85	0.69	0.61	0.74	0.64
13	"4282650"	Little Otter	147.63	0.81	0.77	0.71	0.8
14	"4282780"	Lewis Creek	199.43	0.81	0.76	0.77	0.73
15	"4282795"	Laplatte	113.96	0.82	0.81	0.72	0.77
16	"4290500"	Winooski	2,703.95	0.87	0.83	0.81	0.77
17	"4292500"	Lamoille	1,776.73	0.88	0.83	0.86	0.83
18	"4294000"	Missisquoi	2,201.49	0.87	0.84	0.83	0.81
			Mean	0.83	0.73	0.75	0.72

Note. The drainage areas correspond to those of HYDROTEL. The numbers associated to the basins correspond to them seen on Figure 5.

Table 3
Performance of HYDROTEL Using the KGE Comparing the Local + Global and Global Model Configurations Over the Calibration and Validation Periods for Sum18 and Richelieu River Flow (Fryers)

	Local + global		Global	
	Calibration (1992–2003)	Validation (2004–2013)	Calibration (1992–2003)	Validation (2004–2013)
Sum18	0.93	0.89	0.92	0.89
Fryers	0.87	0.89	0.87	0.9

period and 0.72 for the validation period. When comparing local and global performances over the validation period, we can essentially conclude that the two parameter sets provide equal performances with respective KGE value of 0.73 and 0.72 despite the much larger number of parameters for the local calibration.

An intuitive solution would be to use the model in a local + global configuration mode where one would apply the local parameters for the gauged basin and the global parameters over ungauged basins. This approach was tested and compared to the results of the more simple global configuration where the global parameters are applied ev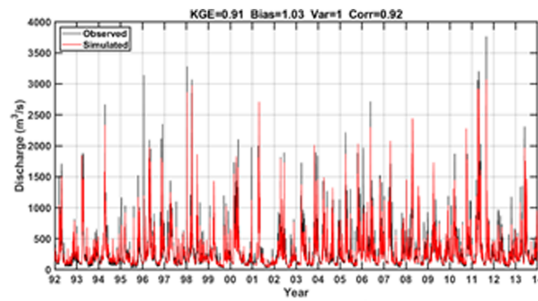erywhere. Table 3 presents the performance indicators for these two configurations. These were compared from a more integrative perspective by using two sets of observations. The first is the sum of the 18 subbasin daily stream flows (hereafter called sum18) which is a good proxy of the total water flow entering the lake, which is of prime interest. The second set of observations consists of the observed Richelieu River flows at the Fryers hydrometric station located downstream of Lake Champlain. Results in Table 3 show that when comparing water entering (sum18) and leaving the lake (Fryers), the two model configurations essentially have equal performances. Surprisingly, the local + global configuration does not seem better than the global configuration in simulating the sum of the 18 subbasins despite showing individually greater KGEs for each of these basins over the calibration period. This suggests that the errors in stream flow simulation that lead to smaller KGE values for the global calibration over the calibration period (see Table 2) compensate each other. It must also be kept in mind that when using sum18 as a basis of comparison, larger weights are given to the larger basins, which typically presents greater KGEs than those of the smaller basins. Results from Tables 2 and 3 tend to demonstrate that the local and global calibrations present similar performances over the validation period, as well as in representing lake inflows and outflows. Because of this, and because the global calibration is much more simple from a parametric perspective, it was decided to continue the analysis with the configuration using only the global parameters in the following sections.

These results suggest that the parameters obtained through a global calibration are more robust to a change in period (Table 2). This conclusion is in line with those of Ricard et al. (2013), Haghnegahdar et al. (2014), and Gaborit et al. (2015). The significant loss in performance between the KGE of the calibration and validation periods when using the parameter values of the local calibration suggests overfitting (Beven, 2006). This would also explain why the optimal parameter values obtained through local calibration for each basin show little spatial coherence (not shown). While local heterogeneity in the parameters might not hinder the ability of the model to reproduce stream flows, it is likely to introduce spatial discontinuity in some of the state variables of the model.
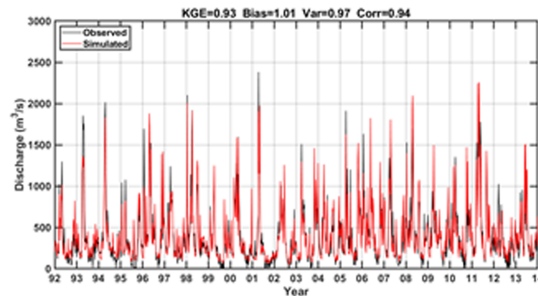
3.2. Assessment of Model Performance for the Simulation of the Multiyear Mean Annual Hydrograph and the 2011 Flood Time Series

To assess model performance, a long hydrological simulation (1992–2013) forced with the Livneh et al. (2015) data set was performed. The simulation was compared with observed stream flows using three independent estimates. First, the HYDROTEL simulation is compared with the observations for the sum of the 18 subbasins (Figure 8). Paying attention to the 1992–2013 time series of the sum18 (top left panel in Figure 8), a quasi-regular pattern with a maximum in spring and minimum in summer can be seen. High flows can sometimes be observed during fall due to heavy precipitation or in winter during warm spells. The KGE value of 0.91 between the simulated and observed sum18 reflects a good simulation of the river flow of the 18 subbasins taken into account using the parameters of the global calibration. On an annual basis, the interannual anomalies of the sum18 are also well reproduced by HYDROTEL from 1992 to 2013 with a KGE value of 0.95 corresponding to a +3% bias of the simulation (top right panel in Figure 8). The simulated 1992–2013 multiyear mean annual hydrograph for the sum18 is close to the observations with a KGE value of 0.93 (top left panel in Figure 9). In this figure, the freshet period, where the flow normally peaks in April, can be clearly seen with a mean inflow of about four to five times that in summer. The large and continuous lake inflow during the months of March, April, and May 2011, which led to the flood, is clearly displayed in the top right panel of Figure 9. Moreover, the very intense, but short duration inflow, at the
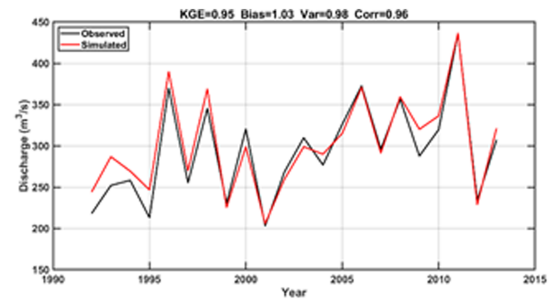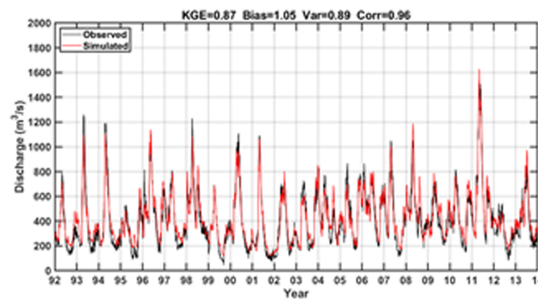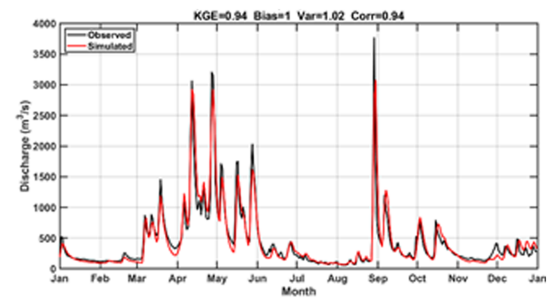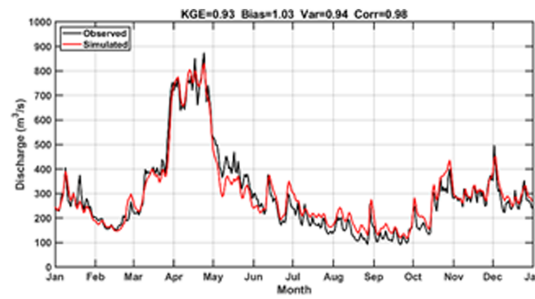
Figure 8. 1992–2013 daily (left column) and annual (right column) time series of the sum of the river flow of the 18 gauged Lake Champlain subbasins (top row), the net basin supply of Lake Champlain (center row), and the Richelieu river flow at Fryers station (bottom row). The measured (black) flow is compared with that simulated by HYDROTEL (red) using the global calibrated parameters and driven with gridded observations.

beginning of September 2011, is associated to the passage of tropical storm Irene. For the sum18, the KGE value of 0.94 was deemed very good, where the overall bias was zero for year 2011.

Second, the HYDROTEL simulation is compared to an estimation of the Lake Champlain NBS from Environment and Climate Change Canada (ECCC; Boudreau et al., 2018) using a quarter of a month time step (center row in Figure 8). The NBS includes the inflows from all the rivers connecting to the Lake Champlain, the precipitation over the lake, and the evaporation of the lake. Again, there is a good match between the simulated and ECCC estimated NBS from 1992 to 2013 at a quarter of a month time step with a KGE value of 0.93. As it was the case with the sum18, the interannual anomalies are well simulated with a KGE value of 0.93 and a bias of +1%. For the 1992–2013 multiyear mean annual and 2011 hydrographs in Figure 9 (center row), the simulated NBS is close to the observed NBS with a KGE value of 0.84 for the annual hydrograph and 0.94 in 2011. The simulated multiyear mean annual hydrograph of NBSs shows a slight underestimation during the mean high flow period in April and May, while the mean low flow is slightly overestimated in August and September. In Figure 9, the simulated peaks of the daily sum18 in 2011 are underestimated compared to the observations, but smoothed over the NBS quarter of a month for, the simulated peaks are comparable to those estimated (right panel).

Third, the simulated and observed Richelieu River flows, downstream of Lake Champlain, at Fryers station are compared. The KGE value of the 1992–2013 daily time series (bottom left panel of Figure 8) is slightly smaller than those for the sum18 and the NBS with a value of 0.87 and a bias of +5%. Most of the positive bias seems to be induced by an overestimation of the summer and fall flows in the model. Despite that,
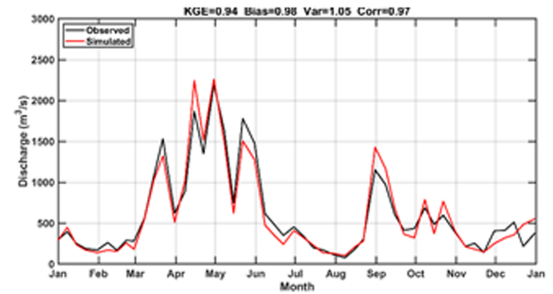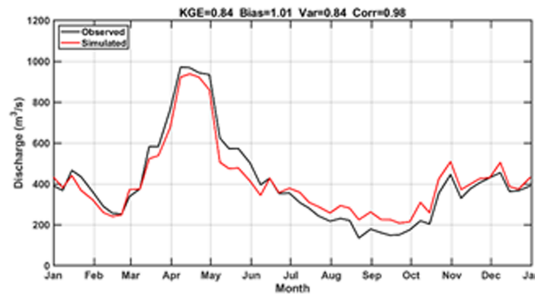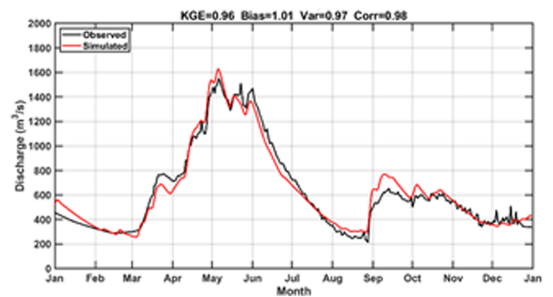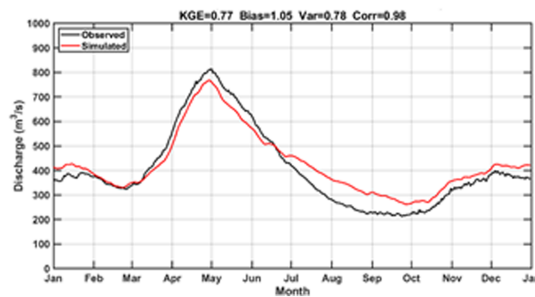
Figure 9. 1992–2013 mean annual hydrograph (left column) and the 2011 time series (right column) of the sum of the river flow of the 18 gauged Lake Champlain subbasins (top row), the net basin supply of Lake Champlain (center row), and the Richelieu river flow at Fryers station (bottom row). The measured (black) flow is compared with that simulated by HYDROTEL (red) using the global calibrated parameters and driven with gridded observations.

the interannual anomalies of the annual means are still well simulated by HYDROTEL with a KGE value of 0.93 (bottom right panel in Figure 8). Looking at the multiyear mean annual hydrograph at Fryers station (bottom left panel in Figure 9), the mean freshet in April and May is underestimated by approximately 5% (50 m^3/s) in HYDROTEL and the mean late summer low flow in August and September is overestimated by 25% (50 m^3/s). Despite those differences, the KGE value of 0.77 for the multiyear mean annual hydrograph and 0.96 in 2011 (bottom right panel in Figure 9) are still acceptable and generally viewed as good. In 2011, HYDROTEL slightly overestimated the observed peak flow of 1,550 m^3/s by 80 m^3/s in early May. The simulation of the river flow at Fryers station remains challenging due to the lake storage and routing effect that are simulated through the adjustment of parameters of the lake's flow rating curve. In addition, the observed river flow is affected by wind surge over the long North-South lake fetch of about 200 km that is not simulated and that can affect the river flow to a certain extent on short timescales (Riboust & Brissette, 2015). Finally, on top of the aforementioned errors, additional errors might come from the hydrological model, the gridded meteorological forcing, and the simulated flow of the Canadian tributaries of the Richelieu River downstream of the lake outlet that were not calibrated. Sources of uncertainties will be further discussed in section 4.

3.3. Analysis of SWE

In the midlatitudes and mountainous regions, a typical hydrograph shows high flows in spring that are mainly associated to the melt of the snowpack that accumulated over winter. Thus, an accurate simulation of the freshet relies on the ability of the hydrological model to simulate the accumulation and melt of the snow. In order to validate the snowpack simulated by HYDROTEL, the snow water equivalent (SWE) of
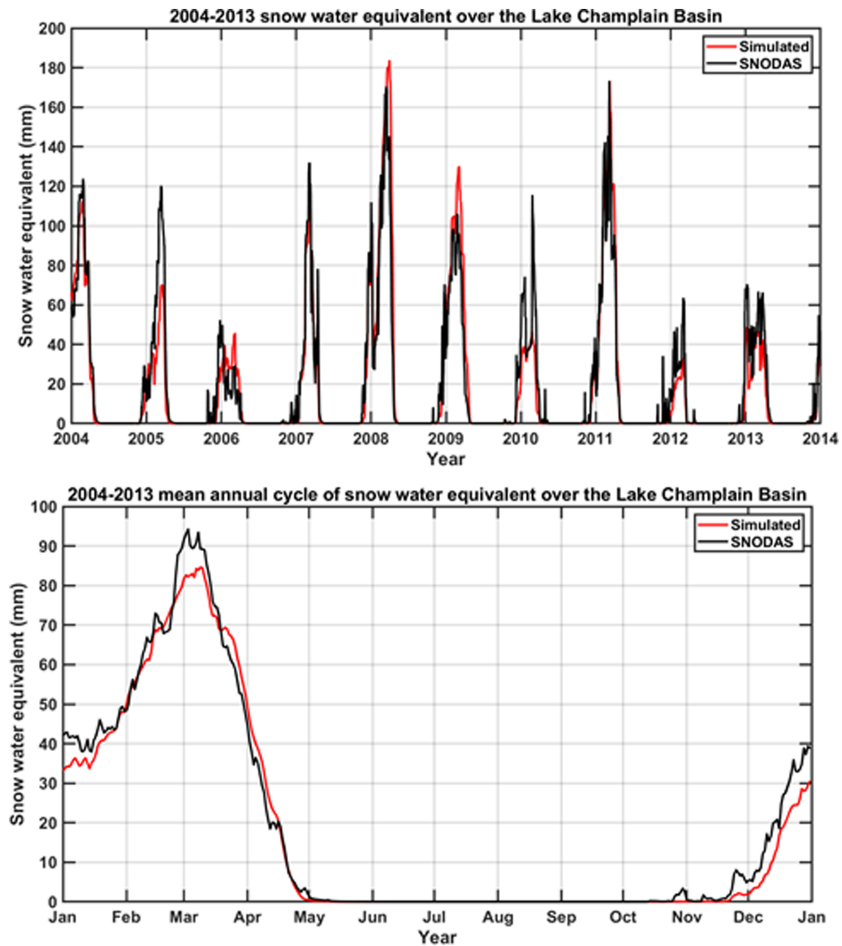
Figure 10. 2004–2013 daily time series (top) and mean annual cycle (bottom) of snow water equivalent simulated by HYDROTEL (red) using the global parameters and driven with gridded observations compared to SNODAS data (black) spatially averaged over the Lake Champlain basin.

the U.S. National Weather Service's Snow Data Assimilation (SNODAS; National Operational Hydrologic Remote Sensing Center, 2004) from the National Oceanic and Atmospheric Organization (NOAA) was used. Available since September 2003, SNODAS is produced through the assimilation of observations from satellite, airborne, and ground-based snow observations into a mathematical model that is driven by a weather prediction model (Carroll et al., 2006). Many variables such as SWE, snow depth, snowmelt, and sublimation are available from the SNODAS on a 1-km gridded data set. In an evaluation of SNODAS for the Colorado Rocky Mountains, Clow et al. (2012) mentioned that SNODAS can provide reliable data as input to moderate-scale to large-scale hydrologic models, which are essential for creating accurate runoff forecasts. The comparison of the Lake Champlain basin spatially averaged simulated SWE by HYDROTEL and that from SNODAS is shown in Figure 10. The successive simulated snowpacks for the 2004–2013 period are in good agreement with SNODAS, as can be seen in the top panel. The 2004–2013 mean annual cycle of SWE (bottom panel in Figure 10) indicates that the mean simulated SWE is slightly underestimated in December (20%) as well as the peak of the snowpack in early-March (10%), but the simulated SWE during the melt from mid-March until the end of April is in very good agreement with SNODAS. SWE over the Richelieu basin peaks at the beginning of March with a basin-averaged around 90 mm. The substantial snowpacks in 2008 and 2011 that led to floods are well simulated by HYDROTEL, but the simulated SWE in 2005, 2010, and 2012 is underestimated. Despite those differences, the simulation of snow by HYDROTEL is satisfying.
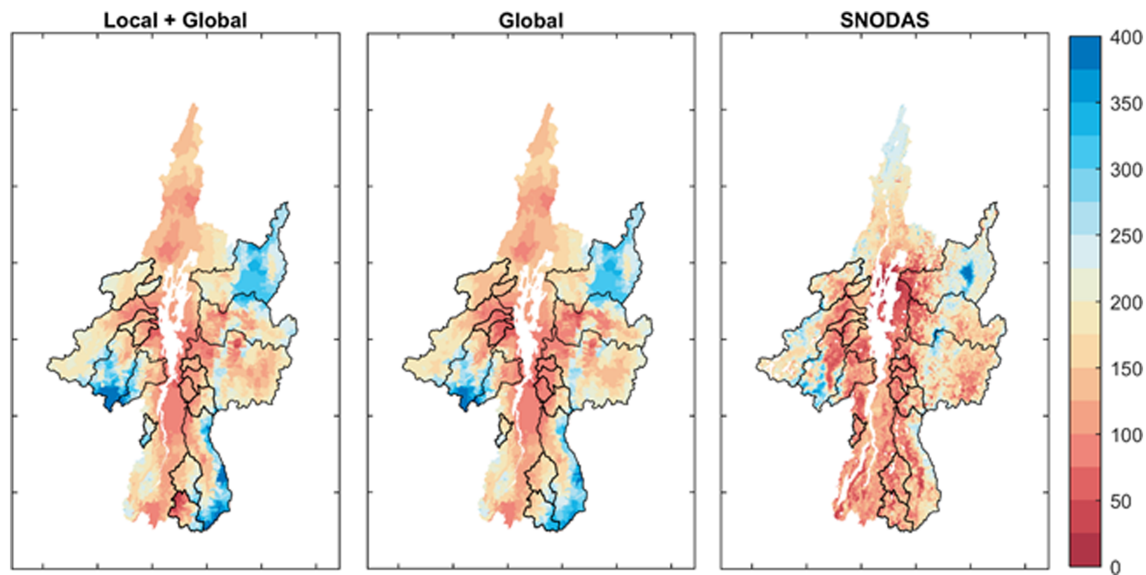
Figure 11. Snow water equivalent (mm) simulated by HYDROTEL on 11 March 2011 using the two approaches (local + global and global parameters only). The values of SWE from SNODAS on the same date are added for comparison.

Figure 11 shows the spatial distribution of SWE over the basin on 11 March 2011 for the two configurations (local + global and global) and the corresponding observationally based gridded values of SNODAS. This date corresponds to when the snowpack reached its maximum prior to the 2011 freshet. We can see that the snowpack distribution is highly heterogeneous across the basin with maximum values reaching over 400 mm in mountainous areas and lower values of 50–100 mm in the valley. The general spatial distribution of SWE of the two HYDROTEL simulations is in agreement with that of SNODAS with small values in the valley and large values in the mountains, but the locations of the peak SWE values in HYDROTEL differ from those of SNODAS.

In the local + global model configuration, some spatial discontinuity can be seen at the boundary of some basins. This is particularly visible in the case of the Mettawee and Poultney basins (#9 and #10 on Figure 5). This is essentially due to the difference in parameters of the model controlling the evolution of the snowpack in the model. These discontinuities might appear inappropriate if one seeks to obtain spatially coherent maps. It must however be kept in mind that none of these maps represent the truth, and they might lead to equally good performance if the main concern is to simulate stream flows. These differences might give a general impression of the uncertainty in the value of the snowpack. An adequate description of this uncertainty would however require to consider uncertainty not only in the model structure parameters but also in the forcing data.

4. Discussion

This study details the application of a high-resolution distributed hydrological model over the Richelieu basin and the ability of this setup to reproduce the streamflow of the Richelieu River (Fryers station) and tributaries of Lake Champlain. In this section, some important points regarding challenges in simulating flows of the Richelieu basin's rivers and related uncertainties are discussed.

4.1. Complexity of the Richelieu Basin

As mentioned above, the Richelieu basin is complex to model because of the rugged terrain of the Adirondack and Green Mountains and the deep valley where Lake Champlain is located. As seen in Figure 6, large regions above 1,000 m of elevation have a very different climate (Figure 3) than those of regions at lower elevations close to the lake. The contrasting temperature between the surroundings of Lake Champlain and the high hills is around 5–6 °C (Figure 3), while precipitation on the windward side of the mountains can be about twice that in the valley because of orographic effects (Figure 3). Due to low
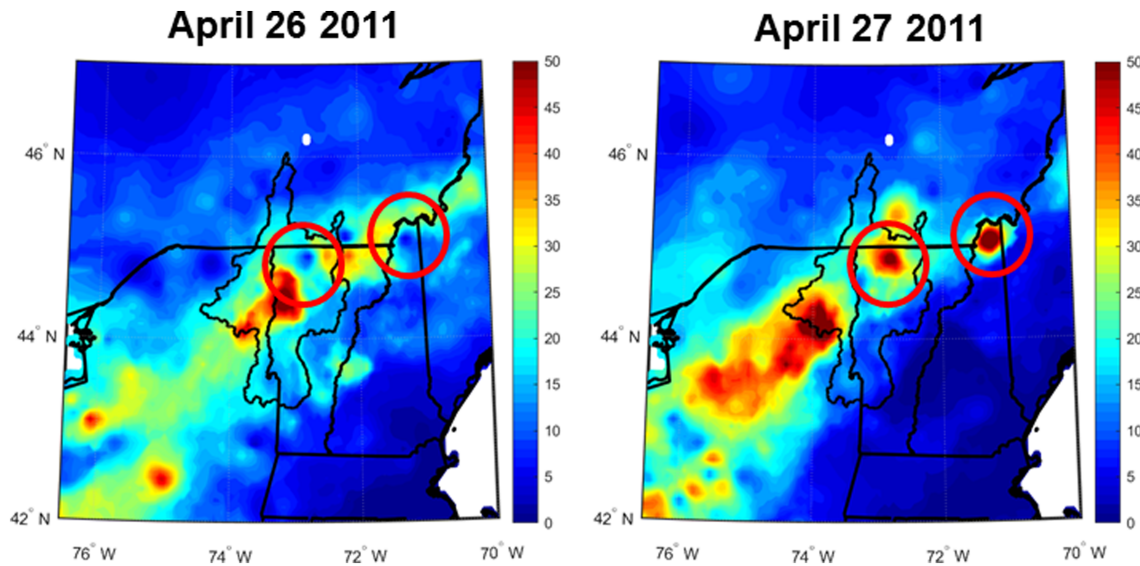
Figure 12. Precipitation (mm/day) from the meteorological data set of Livneh et al. (2015) on 26 and 27 April 2011. Red circles are highlighting unrealistic features.

population density and steep terrain, the density of weather monitoring stations is quite sparse. Thus, the quality and spatial distribution of meteorological data (minimum and maximum temperatures, precipitation) is low and relies largely on interpolation methods. Despite some progress in the Livneh et al. (2015) data set, especially in the reduction of the U.S.-Canada transboundary discontinuities, some problems remain. For example, Figure 12 shows the precipitation on 26 and 27 April 2011. As highlighted in the figure, precipitation amounts are questionable during at least these specific days in northern New Hampshire and Vermont. The minimum values of precipitation on 26 April and the maximum on 27 April in these regions do not match the surrounding precipitation pattern. This particular case could perhaps be explained by a temporal lag in some of the weather stations. This example serves as a reminder that forcing data are not exempt of errors nor are the streamflow observations used as a target to calibrate the model. Inherently, such errors affect the behavior of the model when using the optimization algorithm to identify the parameters during the calibration or by reducing the performance during validation. This probably played a role in the assumed improved robustness of the global calibration. By pooling information together, the impact of these errors is likely limited.

Moreover, due to the rugged terrain in the mountains, most of Lake Champlain's inlets are high-gradient streams, which peak within 24 hr in response to precipitation or snowmelt (Shanley & Denner, 1999). Still, in the mountains, a high percentage of the winter precipitation—stored in the snowpack—melts in spring and reaches Lake Champlain. Thus, the dominant hydrologic event of the year is the spring snowmelt where half of the annual streamflow typically occurs over 2 months (Shanley & Denner, 1999). Because of the geometry of Lake Champlain's outlet, the lake outflow cannot increase at the same rate as the inflows during the spring freshet. Thus, the water level of the lake increases in spring, and the lake acts as a reservoir, attenuating short-term fluctuations of upstream rivers. Consequently, the peak Lake Champlain level lags the peak lake inflows by several days. This latter process, called lake storage and routing, is difficult to simulate because it is affected by many variables. Based on different data sets, Boudreau et al. (2018) established an empirical relationship between the level of the Lake Champlain and the Richelieu River flow that evolves with time due to changes, as instance, of the riverbed vegetation. However, in HYDROTEL, the lake model is rather simple, and the water storage capacity of Lake Champlain does not vary according to the lake level. Thus, the ability of the model to reproduce the lake storage and routing relies on the calibration of the parameters C and k of equation 2. Similarly to the studies of Boyer et al. (2010) and Riboust and Brissette (2015), the current hydrological setup overestimates mean low flows and underestimates mean high flows at the Fryers hydrometric station, which is located downstream of the Lake Champlain outlet. This problem is likely associated to an overestimation of the lake storage and routing and could be improved at the expense of a more targeted calibration. Despite these issues, the overestimation of 50 m^3/s (25%) of the mean low flow

in late summer and the underestimation of 50 m^3/s (5%) of the mean peak flow in spring at Fryers station are rather small and do not affect the performance of the model over the Richelieu basin in an important manner.

4.2. Impact of Model Structure and the Choice of Submodels

Poulin et al. (2011) investigated the effects of model structure and parameter equifinality using a lumped conceptual model (HSAMI) and HYDROTEL over a snow-dominated watershed in southern Quebec. They found that the effect of model structure is more important than that of parameter uncertainty. Even though distributed hydrological models are usually more physically based and likely better in reproducing stream flows in contrasted catchments than lumped models, they heavily rely on physiographic data such as soil type and land cover that may contain errors. For instance, initial calibrations of HYDROTEL over the Richelieu basin led to poor results over the large Otter Creek basin (1643 km^2) when global parameters were used. Further investigations revealed that this basin contains many wetlands that were not taken into account. A custom wetland module was activated in HYDROTEL to account for their presence over the basin. The wetland module is described in details in Fossey et al. (2015). The default wetland parameter values were used over the whole Richelieu River basin except for the Otter Creek basin. For the latter, a manual adjustment of the isolated wetlands was done by adjusting the normal and maximum values of the water depth to 0.5 and 1 m, respectively. Moreover, the parameter controlling the surface occupied by the water in the wetland was set to 1. These parameters were assumed to be homogeneous upstream of the Otter Creek stream gauge. Consequently, the performance of the model using the parameters of the global calibration over this basin significantly improved from a KGE value of 0.57 to 0.81 recalibrating the model. The fact that the simulated streamflow can be improved by providing a better description of the water routing throughout the network (i.e., wetland, river, or lake routing) without having to revisit the hydrological parameters controlling the runoff generation such as those controlling evapotranspiration, snowmelt, and water flow through the soil layer is an interesting advantage of distributed models over lumped ones that should be further investigated.

Moreover, in a previous model setup, the potential evapotranspiration (PET) was simulated with the Hydro-Quebec equation described in Seiller and Anctil (2016), who tested several equations in simulating PET. Another available model setup where the PET is simulated by the Linacre equation, also described in Seiller and Anctil (2016), was then tested. The performance of the simulated multiyear mean annual hydrograph at Fryers station for the 1992–2013 time period significantly improved from a KGE of 0.67 to a KGE of 0.77 when the Linacre equation was used. This result is in agreement with the work of Dallaire et al. (2019), which tested the sensitivity of a lumped model to several PET equations. They found that over the Saumon basin (a neighboring basin of the Richelieu basin), the simulated PET using the Hydro-Quebec equation (based on the difference between daily minimum and maximum temperatures) is among the lowest in summer and highest in winter compared to the simulation of PET with the other equations (Linacre was not used in Dallaire et al., 2019).

4.3. Advantages of Distributed Hydrological Models

Distributed hydrological models such as HYDROTEL are demanding in computing resources to run and usually contain large parameter search spaces, which leads to a more difficult calibration process and longer calibration times. However, for the calibration of HYDROTEL over the Richelieu basin, a new hybrid optimization algorithm specially developed for computationally intensive hydrological model setups was used (Huot et al., 2019). The hybrid approach of this optimization algorithm takes advantage of the benefits of two well-known algorithms (DDS and MADS). The use of a distributed hydrological model (HYDROTEL) and the calibration approach (local calibration followed by a global calibration) are more robust and improve the overall hydrological simulation compared to a lumped model only calibrated at the Fryers hydrometric station. This is because more information is provided to the distributed model with intermediary gauges than in the lumped model with a single outflow. Likewise, the distributed HYDROTEL model also better represents the underlying hydrological processes as compared to a lumped model. Finally, the calibration of the Lake Champlain tributaries ensures a realistic simulation of the streamflow upstream to the lake, something that would not have been possible by just performing a calibration at the basin outlet.

Studies have shown that for many basins, lumped hydrologic models can simulate stream flows that show similar or even better performance than those of more complex distributed or semidistributed model (Lobligeois et al., 2014; Smith et al., 2012). However, in some cases, the use of a distributed model is superior. For instance, Lobligeois et al. (2014) showed that for basin exhibiting highly variable precipitation fields, the use of a distributed over a lumped model improved streamflow simulations. The benefits of distributed models is nevertheless wider than their ability to achieve better streamflow simulations (Tran et al., 2018). They provide spatially distributed information of key hydrological variables and processes. This distributed aspect can become particularly important in the context of impact or mitigation studies (Mauser & Bach, 2009; Paudel et al., 2010). In the present case, the choice of a distributed model is also supported by the fact that it will be used to assess the impact of climate change on the hydrology of the basin and to evaluate potential adaptation solutions to mitigate flood risks.

4.4. Calibration of HYDROTEL

An improved hydrological simulation could have been obtained if the 18 sets of parameters from the local calibration would have been used, but the overall improvement would have been limited by the lake storage and routing parameterization and calibration, which would have required a global calibration in any case. A global set of parameters also has the strong advantage of reducing the number of sets of parameters to carry while ensuring the spatial coherence of the parameters of the same region and allowing the extension of the simulated domain to ungauged neighboring basins. In the current study, calibration was performed over 10 years and validation over the remaining 8 years. The calibration over 10 years was done in order to cover multiple years and preserve an independent period for validation. It would have been preferable to use the complete available period (20 years) to calibrate the model (Arsenault et al., 2018), but this would have been done at the expense of a rather highly computationally intensive calibration while the present case appears as a reasonable trade-off between computational effort and quality of the results.

Possible solutions to reduce overfitting would be (a) to increase the length of calibration period, (b) to reduce the bounds for the calibration of the open parameters, and (c) to decrease the number of open parameters. These solutions aim to reduce the number of degrees of freedom (open parameters) relative to the available information (length of observed records). In the present case, the goal was to keep a sufficiently long period for model validation that, considering the data availability, explains the calibration period of 11 years. However, once the assessment of the robustness is done, studies suggest that one should use all of the available data to calibrate the model prior to use (Arsenault et al., 2018). For the local calibration, a reduction in open parameters could probably have been done through a sensitivity analysis. However, it is likely that different basins would have converged to different sets of parameters, which adds to the complexity of calibration. On the other hand, with the global calibration, such analysis seems unnecessary, as the global calibration scheme shows no indication of overfitting. Nevertheless, it can be presumed that there is probably an ideal parametric complexity that lies somewhere between the local calibration that seems too complex and the global one that seems too simple. For instance, an interesting idea to test intermediate complexity would be to explore the possibility of calibrating some parameters under the assumption that they are homogenous among all subbasins (global assumption) while other parameters would vary locally.

4.5. Additional Hydrological Processes

Despite being distributed and more physically based than lumped models, HYDROTEL does not simulate some important physical processes affecting river flow and lake levels. In particular, as most hydrological models, HYDROTEL does not take into account wind effect on lake levels. For instance, steady wind channeling conditions along the long North-South lake fetch can create surges in lake levels, especially at the north end outlet of Lake Champlain, resulting in an increase/decrease of the Richelieu River flow. Only sophisticated 3D lake models can simulate such wind effects, and those are particularly computationally expensive. Fortunately, steady wind conditions affecting river flows are rare and happen over short timescales when specific weather patterns occur. Thus, the absence of this hydrological process in the simulation is not a major concern for this study but could explain why some simulated flows do not temporarily match the observations. It is also noteworthy that simulating wind conditions at the lake surface or obtaining accurate wind measurements over Lake Champlain remains a challenge.

4.6. Errors in Measurements of River Flow, NBS, and SWE

The calibration process of HYDROTEL is heavily dependent on the quality of the meteorological forcing as well as observed river flows used as calibration target; none of which are free from errors. Most of the time, the stream flows are obtained through measurements of water levels, which are then converted to discharges using a rating curve. In general, rating curves contain many uncertainties related to the modification of the riverbed in time, difficulties in establishing flow velocity profiles, and extrapolation between rating curve values (Domeneghetti et al., 2012). Moreover, for a given discharge, the water level of a river can be affected by the presence of ice, plants, or other debris in the control section. Errors also arise in the computation of the Lake Champlain NBS provided by ECCC for validation. In fact, even though they are plausible, some negative NBS values demonstrate the possibility of errors in the water budget computed through the approximations of the lake outflow in the measured river flow or in the uncertainties in the measure of the lake levels affected by lake seiche (Boudreau et al., 2018). Finally, SWE from SNODAS contains errors either from the snow model or snow measurements affecting the data assimilation. Overall, one should always keep in mind that the calibration and validation of a hydrological model can be affected by errors in the observations of the river flow and alternative variables such as NBS and SWE.

5. Conclusions

This paper presented the different steps leading to the application of an advanced high-resolution distributed hydrological model on the Richelieu River basin. So far, in the literature, the river flow of the Richelieu basin had mainly been simulated with lumped models and in a semidistributed fashion still with a lumped model. While they are lower in complexity and lead to an easy and rapid application, simplified lumped models have shown deficiencies on the target basin by underestimating freshets and overestimating low flows. Due to the presence of steep terrain and a large lake, the application of a hydrological model over the Richelieu basin is particularly challenging. Taking advantage of recent advances in hydrological modeling, the current application of HYDROTEL included several innovations such as high-resolution discretization of the basin, local and global calibrations of tributaries, accounting for river flow and lake inflow, a novel hybrid optimization algorithm, and an up-to-date gridded observed meteorological data. Taking into account these aspects, it was expected that the simulation of the river flow of the Richelieu would be more successful than other studies found in the literature.

To start, the model validation was done through a comparison of the total inflow to Lake Champlain from 18 subbasins (sum18) both observed and simulated. Taking advantage of the calibration, using an objective function weighted according to the surface area of the subbasins, the simulated sum18 showed a close match to that observed with KGE values of 0.93 for the multiyear mean annual hydrograph and 0.94 for 2011 when the record-breaking flood occurred. In a second step, the validation of the simulated Lake Champlain NBS was compared to that estimated by ECCC and showed a good match with KGE values of 0.84 for the multi-year mean annual hydrograph and 0.94 for 2011. Finally, the simulated Richelieu River flow at the Fryers hydrometric station was in good agreement with the observed flow but suffered from an overestimation (25%) of the mean low flows and underestimation (5%) of high flows, but these were seemingly smaller than those reported in the literature. The calibration of the rating curve at the outlet of Lake Champlain (i.e., discharge vs. water level mimicking storage and routing effects) is all likelihood an important contributor of these biases, since the sum18 and the NBS were in very good agreement with observations. The validation of the simulated SWE revealed a realistic simulation of the accumulation of snow and snowmelt by HYDROTEL in a comparison with SWE estimates (SNODAS) from NOAA.

Many errors from either the hydrological model HYDROTEL, the identification of the soil type and land cover, the calibrated parameter values, and the gridded observations, to name a few, could have affected the performance of the hydrological model. Moreover, the measurements of the river flows, the NBS estimates, and the SWE from SNODAS contain errors that can affect the calibration and validation of the model. Overall, the performance of the application of HYDROTEL on the Richelieu River basin was judged satisfactory and adequate to move forward to the next step consisting of forcing HYDROTEL with bias-corrected climate projections. This next step, which focusses on assessing the impacts of climate change on the river flow, particularly on extremes such as floods and low flows, will be presented in a forthcoming paper.

References

Abaza, M., Anctil, F., Fortin, V., & Turcotte, R. (2014). Sequential streamflow assimilation for short-term hydrological ensemble forecasting. *Journal of Hydrology*, *519*, 2692–2706.

Abaza, M., Anctil, F., Fortin, V., & Turcotte, R. (2015). Exploration of sequential streamflow assimilation in snow dominated watersheds. *Advances in Water Resources*, *80*, 79–89.

Arsenault, R., Brissette, F., & Martel, J.-L. (2018). The hazards of split-sample validation in hydrological model calibration. *Journal of Hydrology*, *566*, 346–362. https://doi.org/10.1016/j.jhydrol.2018.09.027

Audet, C., & Dennis, J. E. Jr. (2006). Mesh adaptive direct search algorithms for constrained optimization. *SIAM Journal on Optimization*, *17*, 188–217.

Beven, K. (2006). A manifesto for the equifinality thesis. *Journal of Hydrology*, *30*, 18–36.

Blanchette, M., Rousseau, A. N., Foulon, E., Savary, S., & Poulin, M. (2019). What would have been the impacts of wetlands on low flow support and high flow attenuation under steady state land cover conditions? *Journal of Environmental Management*, *234*, 448–457. https://doi.org/10.1016/j.jenvman.2018.12.095

Bontemps, S., Defourny, P., Van Bogaert, E., Arino, O., Kalogirou, V., & Ramos Perez, J. (2011). GLOBCOVER 2009 Products description and validation report. *Université catholique de Louvain (UCL) & European Space Agency (esa)*, **Vers. 2.2**, 53 pp.

Bouda, M., Rousseau, A. N., Gumiere, S. J., Gagnon, P., Konan, B., & Moussa, R. (2014). Implementation of an automatic calibration procedure for HYDROTEL based on prior OAT sensitivity and complementary identifiability analysis. *Hydrological Processes*, *28*, 3947–3961.

Bouda, M., Rousseau, A. N., Konan, B., Gagnon, P., & Gumiere, S. J. (2012). Case study: Bayesian uncertainty analysis of the distributed hydrological model HYDROTEL. *Journal of Hydrologic Engineering*, *17*, 1021–1032.

Boudreau, P., Morin, J., & Champoux, O. (2018). Modèle de bilan hydrique et apports nets au lac Champlain (Note technique NT-124). Section hydrologie et écohydraulique. Services hydrologiques nationaux. Service Météorologique du Canada, Environnement et Changement Climatique Canada. 26 p.

Boyer, C., Chaumont, D., Chartier, I., & Roy, A. G. (2010). Impact of climate change on the hydrology of St. Lawrence tributaries. *Journal of Hydrology*, *384*(1–2), 65–83.

Carroll, T. R., Cline, D. W., Olheiser, C., Rost, A., Nilsson, A., Fall, G., et al. (2006). NOAA's national snow analyses. Proceedings of the 74th annual meeting of the Western snow conference 74: 13.

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*, W12509. https://doi.org/10.1029/2011WR010602

Clow, D. W., Nanus, L., Verdin, K. L., & Schmidt, J. (2012). Evaluation of SNODAS snow depth and snow water equivalent estimates for the Colorado Rocky Mountains, USA. *Hydrological Processes*, *26*, 2583–2591. https://doi.org/10.1002/hyp.9385

Dallaire, G., Poulin, A., & Arsenault, R. (2019). Investigating the uncertainty of potential evapotranspiration modelling in climate change impact studies on water resources in Canada, United States and Mexico. Geophysical Research Abstracts, Vol. 21, EGU2019–11204, EGU General Assembly 2019.

Direction de l'expertise hydrique (2018a). Document d'accompagnement de l'Atlas hydroclimatique du Québec méridional. Ministère du Développement durable, de l'Environnement et de la Lutte contre les changements climatiques, 34 p. http://www.cehq.gouv.qc.ca/atlas-hydroclimatique/doc-accompagnement.pdf accessed April 1, 2019.

Direction de l'expertise hydrique (2018b). Mise à jour de l'atlas hydroclimatique du Québec meridional, rapport technique, Ministère du Développement durable, de l'Environnement et de la Lutte contre les changements climatiques, 54 p.

Domeneghetti, A., Castellarin, A., & Brath, A. (2012). Assessing rating-curve uncertainty and its effects on hydraulic model calibration. *Hydrology and Earth System Sciences*, *16*, 1191–1202.

Essou, G. R. C., Brissette, F., & Lucas-Picher, P. (2017). The use of reanalyses and gridded observations as weather input data for a hydrological model: Comparison of performances of simulated river flows based on the density of weather stations. *Journal of Hydrometeorology*, *18*, 497–513.

Fortin, J. P., Moussa, R., Bocquillon, C., & Villeneuve, J. P. (1995). HYDROTEL, a distributed hydrological model compatible with remote sensing and geographical information systems. *Journal of Water Science*, *8*, 97–124.

Fortin, J. P., Turcotte, R., Massicotte, S., Moussa, R., & Fitzback, J. (2001). A distributed watershed model compatible with remote sensing and GIS data, part 1: Description of the model. *Journal of Hydrologic Engineering, ASCE*, *6*(2).

Fossey, M., & Rousseau, A. N. (2016a). Can isolated and riparian wetlands mitigate the impact of climate change on watershed hydrology? A case study approach. *Journal of Environmental Management*, *182*, 327–339.

Fossey, M., & Rousseau, A. N. (2016b). Assessing the long-term hydrological services provided by wetlands under changing climate conditions: A case study approach of a Canadian watershed. *Journal of Hydrology*, *541*, 1287–1302.

Fossey, M., Rousseau, A. N., Bensalma, F., Savary, S., & Royer, A. (2015). Integrating isolated and riparian wetland modules in the PHYSITEL/HYDROTEL modelling platform: Model performance and diagnosis. *Hydrological Processes*, *29*, 4683–4702.

Fossey, M., Rousseau, A. N., & Savary, S. (2016). Modelling the hydrological impacts of wetlands in the Becancour River watershed, Canada: A spatio-temporal dependent effect. *Hydrological Processes*, *30*, 1768–1781.

Foulon, É., Rousseau, A. N., & Gagnon, P. (2018). Development of a methodology to assess future trends in low flows at the watershed scale using solely climate data. *Journal of Hydrology*, *557*, 774–790.

Gaborit, É., Ricard, S., Lachance-Cloutier, S., Anctil, F., & Turcotte, R. (2015). Comparing global and local calibration schemes from a differential split-sample test perspective. *Canadian Journal of Earth Sciences*, *52*, 990–999. https://doi.org/10.1139/cjes-2015-0015

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.

Haghnegahdar, A., Tolson, B. A., Davison, B., Seglenieks, F. R., Klyszejko, E., Soulis, E. D., et al. (2014). Calibrating environment Canada's MESH modelling system over the Great Lakes basin. *Atmosphere-Ocean*, *52*(4), 281–293. https://doi.org/10.1080/07055900.2014.939131

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*, *64*, 1204–1222.

International Joint Commission, (2017). Work plan for the identification of measures to mitigate flooding and the impacts of flooding of the Lake Champlain and Richelieu River. https://ijc.org/sites/default/files/2018-07/LCRR_final_work_plan.pdf, .

Kling, H., Fuchs, M., & Paulin, M. (2012). Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios. *Journal of Hydrology*, *424*, 264–277. https://doi.org/10.1016/j.jhydrol.2012.01.001

Lavigne, M.-P., Rousseau, A. N., Turcotte, R., Laroche, A.-M., Fortin, J.-P., & Villeneuve, J.-P. (2004). Validation and use of a distributed hydrological modeling system to predict short term effects of clear cutting on the hydrological regime of a watershed. *Earth Interactions*, *8*(3), 1–19.

Linacre, E. T. (1977). A simple formula for estimating evaporation rates in various climates, using temperature data alone. *Agricultural Meteorology*, *18*, 409–424.

Livneh, B., Bohn, T. J., Pierce, D. W., Munoz-Arriola, F., Nijssen, B., Vose, R., et al. (2015). A spatially comprehensive, hydrometeorological data set for Mexico, the US, and southern Canada 1950–2013. *Scientific Data*, *2*, 150,042. https://doi.org/10.1038/sdata.2015.42

Lobligeois, F., Andréassian, V., Perrin, C., Tabary, P., & Loumagne, C. (2014). When does higher spatial resolution rainfall information improve streamflow simulation? An evaluation using 3620 flood events. *Hydrology and Earth System Sciences*, *18*(2), 575–594. https://doi.org/10.5194/hess-18-575-2014

Lucas-Picher, P., Riboust, P., Somot, S., & Laprise, R. (2015). Reconstruction of the spring 2011 Richelieu river flood by two regional climate models and a hydrological model. *Journal of Hydrometeorology*, *16*, 36–54. https://doi.org/10.1175/JHM-D-14-0116.1

Ludwig, R., May, I., Turcotte, R., Vescovi, L., Braun, M., Cyr, J.-F., et al. (2009). The role of hydrological model complexity and uncertainty in climate change impact assessment. *Advances in Geosciences*, *21*, 63–71. http://www.adv-geosci.net/21/63/2009/

Mauser, W., & Bach, H. (2009). PROMET—Large scale distributed hydrological modelling to study the impact of climate change on the water flows of mountain watersheds. *Journal of Hydrology*, *376*(3), 362–377. https://doi.org/10.1016/j.jhydrol.2009.07.046

Ministère des Ressources naturelles et de la Faune, Gouvernement du Québec. (2008). La Base de données topographiques du Québec à l'échelle de 1/20 000 (BDTQ 20k) https://mern.gouv.qc.ca/publications/territoire/expertise/08_BDTQ_mai2008.pdf

National Operational Hydrologic Remote Sensing Center (2004). Snow Data Assimilation System (SNODAS) Data Products at NSIDC, Version 1. (Masked Data (over United States Only)). Boulder, Colorado USA. NSIDC: National Snow and Ice Data Center. doi: https://doi.org/10.7265/N5TB14TC. (Accessed in December 2018).

Noël, P., Rousseau, A. N., Paniconi, C., & Nadeau, D. F. (2014). An algorithm for delineating and extracting hillslopes and hillslope width functions from gridded elevation data. *Journal of Hydrologic Engineering*, *19*(2), 366–374.

Oreiller, M., Nadeau, D. F., Minville, M., & Rousseau, A. N. (2014). Modelling snow water equivalent and spring runoff in a boreal watershed, James Bay, Canada. *Hydrological Processes*, *28*, 5991–6005.

Paudel, M., Nelson, E. J., Downer, C. W., & Hotchkiss, R. (2010). Comparing the capability of distributed and lumped hydrologic models for analyzing the effects of land use change. *Journal of Hydroinformatics*, *13*(3), 461–473. https://doi.org/10.2166/hydro.2010.100

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.

Riboust, P., & Brissette, F. (2015). Climate change impacts and uncertainties on spring flooding of Lake Champlain and the Richelieu River. *Journal of the American Water Resources Association*, *51*, 776–793.

Riboust, P., & Brissette, F. (2016). Analysis of Lake-Champlain/Richelieu River's historical 2011 flood. *Canadian Water Resources Association Journal*, *41*, 174–185.

Ricard, S., Bourdillon, R., Roussel, D., & Turcotte, R. (2013). Global calibration of distributed hydrological models for large scale applications. *Journal of Hydrological Engineering*, *18*(6), 719–721. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000665

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*, *41*, 129–138. https://doi.org/10.1080/07011784.2014.999825

Samuel, J., Rousseau, A. N., Abbasnezhadi, K., & Savary, S. (2019). Development and evaluation of a hydrologic data-assimilation scheme for short-range flow and inflow forecasts in a data-sparse high-latitude region using a distributed model and ensemble Kalman filtering. *Advances in Water Resources*, *130*, 198–220. https://doi.org/10.1016/j.advwatres.2019.06.004

Seiller, G., & Anctil, F. (2016). How do potential evapotranspiration formulas influence hydrological projections. *Hydrological Sciences Journal*, *61*, 2249–2266.

Shangguan, W., Dai, Y., Duan, Q., Liu, B., & Yuan, H. A. (2014). A global soil data set for earth system modeling. *Journal of Advances in Modeling Earth Systems*, *6*, 249–263. https://doi.org/10.1002/2013MS000293

Shanley, J. B., & Denner, J. C. (1999). The hydrology of the Lake Champlain Basin. In *Lake Champlain in transition: From research toward restoration*, (pp. 41–66). Washington, D.C: American Geophysical Union.

Smith, M. B., Koren, V., Zhang, Z., Zhang, Y., Reed, S. M., Cui, Z., et al., & DMIP 2 Participants (2012). Results of the DMIP 2 Oklahoma experiments. *Journal of Hydrology*, *418–419*, 17–48. https://doi.org/10.1016/j.jhydrol.2011.08.056

Tolson, B. A., & Shoemaker, C. A. (2007). Dynamically dimensioned search algorithm for computationally efficient watershed model calibration. *Water Resources Research*, *43*, W01413. https://doi.org/10.1029/2005WR004723

Tran, Q. Q., De Niel, J., & Willems, P. (2018). Spatially distributed conceptual hydrological model building: A generic top-down approach starting from lumped models. *Water Resources Research*, *54*, 8064–8085. https://doi.org/10.1029/2018WR023566

Turcotte, R., Fortin, J.-P., Rousseau, A. N., Massicotte, S., & Villeneuve, J.-P. (2001). Determination of the drainage structure of a watershed using a digital elevation model and a digital river and lake network. *Journal of Hydrology*, *240*, 225–242.

Turcotte, R., Fortin, L. G., Fortin, V., Fortin, J. P., & Villeneuve, J. P. (2007). Operational analysis of the spatial distribution and the temporal evolution of the snowpack water equivalent in southern Quebec. *Canadian Nordic Hydrology*, *38*(3), 211–234.

Turcotte, R., Lacombe, P., Dimnik, C., & Villeneuve, J. P. (2004). Distributed hydrological prediction for the management of Quebec's public dams/prevision hydrologique distribuée pour la gestion des barrages publics du Quebec. *Canadian Journal of Civil Engineering*, *31*(2), 308–320.

Turcotte, R., Rousseau, A. N., Fortin, J.-P., & Villeneuve, J.-P. (2003). Development of a process-oriented, multiple-objective, hydrological calibration strategy accounting for model structure. In Q. Duan, H. V. Gupta, S. Sorooshian, A. N. Rousseau, & R. Turcotte (Eds.), *Advances in Calibration of Watershed Models, Water Science & Application*, (Vol. 6, pp. 153–163). Washington, USA: American Geophysical Union (AGU).

United Geological of the United States (1999). National elevation dataset. https://pubs.usgs.gov/fs/1999/0148/report.pdf

Velazquez, J. A., Schmid, J., Ricard, S., Muerth, M. J., Gauvin-St-Denis, B., Minville, M., et al. (2013). An ensemble approach to assess hydrological models' contribution to uncertainties in the analysis of climate change impact on water resources. *Hydrology and Earth System Sciences*, *17*(2), 565–578. https://doi.org/10.5194/hess-17-565-2013