Enhancing physically based and distributed hydrological model calibration through internal state variable constraints

Frédéric Talbot[1], Jean-Daniel Sylvain [1,2], Guillaume Drolet[2], Annie Poulin[1], Richard Arsenault[1]

[1] Hydrology, Climate and Climate Change Laboratory, École de technologie supérieure, Université du Québec, Montréal, H3C 1K3, Canada

[2] Direction de la recherche forestière, Ministère des Ressources naturelles et des Forêts, Québec, G1P 3W8, Canada

Correspondence to: Frédéric Talbot (frederic.talbot.2@ens.etsmtl.ca)

Abstract. Accurately representing hydrological processes remains a major challenge in hydrological modeling. Recent studies have demonstrated the benefits of multi-variable calibration, which integrates additional hydrological variables such as evapotranspiration and soil moisture alongside streamflow to improve model realism. However, groundwater recharge as a calibration variable remains relatively underexplored.

This study evaluates how incorporating groundwater recharge into the calibration of the Water Balance Simulation Model (WaSiM) affects hydrological variables representation. Three configurations were tested: Baseline (BL) with streamflow-only calibration, Physical Groundwater Model (GW) with physically-based groundwater flow, and Physical Groundwater with Recharge Calibration (GW-RC), which further constrains groundwater recharge during calibration. The models were calibrated and applied to 34 catchments in Southern Québec. Their performance was evaluated using the Kling-Gupta Efficiency (KGE) for streamflow and spatial estimates of groundwater recharge derived from a previous research project conducted in the same region.

Results indicate that while calibrating on streamflow alone produces high KGE values (median KGE = 0.83 for GW and 0.82 for BL), but it comes at the cost of misrepresenting subsurface hydrological processes. Adding groundwater recharge constraints (GW-RC) reduce streamflow performance, with a median KGE of 0.77 for GW-RC, but improves hydrological variable representation, especially in seasonal runoff patterns, where it better captures the balance between surface runoff and interflow during snowmelt. Additionally, GW-RC showed the smallest differences with the groundwater recharge estimates.

These findings illustrate the consequence of equifinality in streamflow-based calibration, where multiple parameter sets can yield similar streamflow outputs while misrepresenting internal hydrological processes. Incorporating groundwater recharge constraints improves the representation of internal hydrological processes while maintaining strong streamflow simulation performance, which could ultimately enhance reliability of climate change adaptation and water resource management strategies.

1 Introduction

Accurately representing watershed processes under climate change remains a central challenge in the evolving field of hydrology (Persaud et al., 2020). Recent advances in hydrological modeling have offered valuable insights into water resource management and climate adaptation strategies (Xu et al., 2005; Chen et al., 2011; Wang et al., 2023). However, the complexity of watershed dynamics, especially in snow dominated catchments, necessitates models that can accurately simulate both surface and subsurface hydrological processes (Chu and Shirmohammadi, 2004; Farjad et al., 2016).

The need for detailed, physically based hydrological modeling goes beyond immediate concerns of water management and climate impact assessments. Groundwater dynamics are crucial for forest health (Maitre et al., 1999; Jacobs, 2003), as stable water availability supports ecosystem resilience (Cunningham et al., 2011; Orellana et al., 2012). By enhancing the accuracy of groundwater simulation and recharge calibration, we can improve our ability to forecast forest growth and resilience under changing climatic conditions (Ford et al., 2011; Grant et al., 2013). This linkage underscores the importance of detailed hydrological modeling and aligns with broader environmental, economic, and ecological management goals aimed at sustaining forest productivity in the face of environmental change. This approach helps forest managers make informed decisions, supporting the long-term health and sustainability of forest ecosystems (Vose et al., 2011; Sun et al., 2023).

The Water balance Simulation Model (WaSiM) (Schulla, 2021) is a distributed and physically based hydrological model. It stands out for its complexity, fine spatial resolution and comprehensive approach to modeling hydrological processes. This makes the model especially useful for analyzing intermediate hydrological variables with greater reliability. Several studies exemplify the application of WaSiM for examining internal hydrological variables across diverse geographic settings and scenarios. For example, Jasper et al. (2006) analyzed summer soil water pattern shifts due to climatic changes, demonstrating that WaSiM could effectively model the substantial alterations in hydrological responses to varying climate scenarios. Natkhin et al. (2012) used WaSiM to differentiate the impacts of climate change and forest growth dynamics on groundwater recharge in Northeast Germany. Similarly, two separate studies (Rößler and Löffler, 2010; Rössler et al., 2012) analyzed soil moisture dynamics using WaSiM, discussing the modeling potentials and limitations in high mountain catchments and the broader impact of climate on soil moisture. Bormann and Elfert (2010) investigated how land use changes influence various runoff generation processes such as surface runoff, interflow, and baseflow. Furthermore, Förster et al. (2017, 2018) conducted detailed comparisons of internal state variables with actual forest measurements, including meteorological variables and snow cover dynamics, highlighting the refined capabilities of WaSiM to model complex interactions like snow cover and canopy interception. These studies collectively demonstrate the model's utility in capturing a wide range of hydrological variables.

Despite recent advances, hydrological modeling still faces challenges in representing watershed dynamics. These challenges are especially evident when calibration relies only on streamflow data (Mei *et al.*, 2023; Schäfer *et al.*, 2023; de Lima Ferreira and da Paz, 2024; Pool *et al.*, 2024). While streamflow is a key indicator for capturing temporal fluctuations in water systems, it offers limited insights into the internal hydrological processes (Rajib *et al.*, 2018). This reliance on

streamflow can result in models that perform well in reproducing observed flows but misrepresent underlying processes. This phenomenon, known as equifinality, occurs when different parameter sets produce the same outputs but for the wrong reasons (Kirchner, 2006; Yassin *et al.*, 2017; Acero Triana *et al.*, 2019; Mei *et al.*, 2023). Therefore, focusing only on streamflow in model calibration can hide important differences in how hydrological processes are represented.

In pursuit of better representing hydrological processes at the catchment scale, several studies have explored hydrologic scaling and parameter transferability (Samaniego *et al.*, 2010, 2017; Mizukami *et al.*, 2017; Imhoff *et al.*, 2020). Notably, Samaniego *et al.* (2010) introduced the multiscale parameter regionalization to tackle overparameterization and the non-transferability of parameters across different scales. Ficchì et al. (2019) also proposed a model structure that considers flow accuracy and fluxes match on different modelling timesteps, adjusting the structure and parameters to ensure robust simulation across various time scales. Additionally, Peters-Lidard *et al.* (2017) advocated for adopting the fourth paradigm of data-intensive science in hydrology, which leverages emerging datasets to refine our understanding of hydrological models and processes. This paradigm suggests that advancements in computational science represent a new methodological branch alongside empiricism, theory, and computational simulation. By enabling the intensive use of data, these advancements can revolutionize science by facilitating the discovery and testing of theories and models. This approach emphasizes the integration of comprehensive datasets and computational tools into conventional scientific workflows, thereby enhancing the capacity for scientific innovation and synthesis in hydrology.

Recent studies have advocated for a shift towards integrating additional hydrological variables and data sources, such as remote sensing products and in-situ measurements, into the calibration process (Dembélé *et al.*, 2020; Meyer Oliveira *et al.*, 2021; Liu *et al.*, 2022; Mei *et al.*, 2023; Schäfer *et al.*, 2023; de Lima Ferreira and da Paz, 2024; Pool *et al.*, 2024). Mei *et al.* (2023) found that including gridded soil moisture alongside gauged streamflow improved evapotranspiration simulations across 20 catchments in the Lake Michigan watershed. Schäfer *et al.* (2023) used WaSiM to simulate the water balance of a forested catchment in Germany, showing that including plant-available water and evapotranspiration data significantly enhanced model accuracy. De Lima Ferreira and da Paz (2024) similarly improved model performance by incorporating actual evapotranspiration estimates into a hydrological model of a Brazilian semi-arid basin, highlighting the benefits of multi-variable calibration and the need to test distinct data sources.

Although many studies have successfully used variables such as soil moisture, evapotranspiration, and groundwater head in model calibration, there remains a gap in understanding how other variables, like groundwater recharge, can improve the representation of hydrological processes. Addressing this gap is important for both the theoretical advancement of hydrological sciences and the practical applications of water resource management, flood risk assessment, and climate change mitigation (Pradhan and Indu, 2019). By adopting a calibration approach that integrates a more holistic view of watershed processes, models become more reflective of complex hydrological interactions and gain robustness in the face of non-stationary climate conditions (Wang *et al.*, 2023). This enhanced process representation and strengthens confidence in model projections, making them more reliable for future applications.

In this study, we implement three distinct model configurations of the WaSiM hydrological model: Baseline (BL), which follows a traditional streamflow-based calibration; Physical Groundwater Model (GW), which introduces physically based groundwater flow processes; and Physical Groundwater with Recharge Calibration (GW-RC), which further constrains groundwater recharge during calibration. The objective is to investigate how different calibration strategies and levels of model complexity influence the representation of hydrological processes over a set of 34 catchments in snowy catchment conditions. Through comparative analysis of these configurations, we aim to expose the nuances in model performance and hydrological variable representation, contributing to the ongoing debate on the best practices for hydrological model calibration.

2 Methods

2.1 Study area

This study examines 34 catchments in Southern Quebec, Canada, each with distinct physiographic and hydrometeorological features. The catchments range in size from 525 to 6,840 km² (see Fig. 1). These catchments were selected based on several key criteria to ensure robust model calibration and validation. Specifically, they were selected based on the availability of comprehensive streamflow data from 1981 to 2010. Additionally, catchments were selected to represent the region's geographical and hydrological diversity to capture a range of climatic conditions across the study area. Where possible, catchments covered by the PACES project (see detail in section 2.2.5) were prioritized to ensure data consistency and facilitate comparisons of groundwater recharge estimates. To preserve the natural integrity of hydrological processes under study, selected catchments needed to be free from dams and reservoirs and located away from major urban areas to minimize anthropogenic influences.

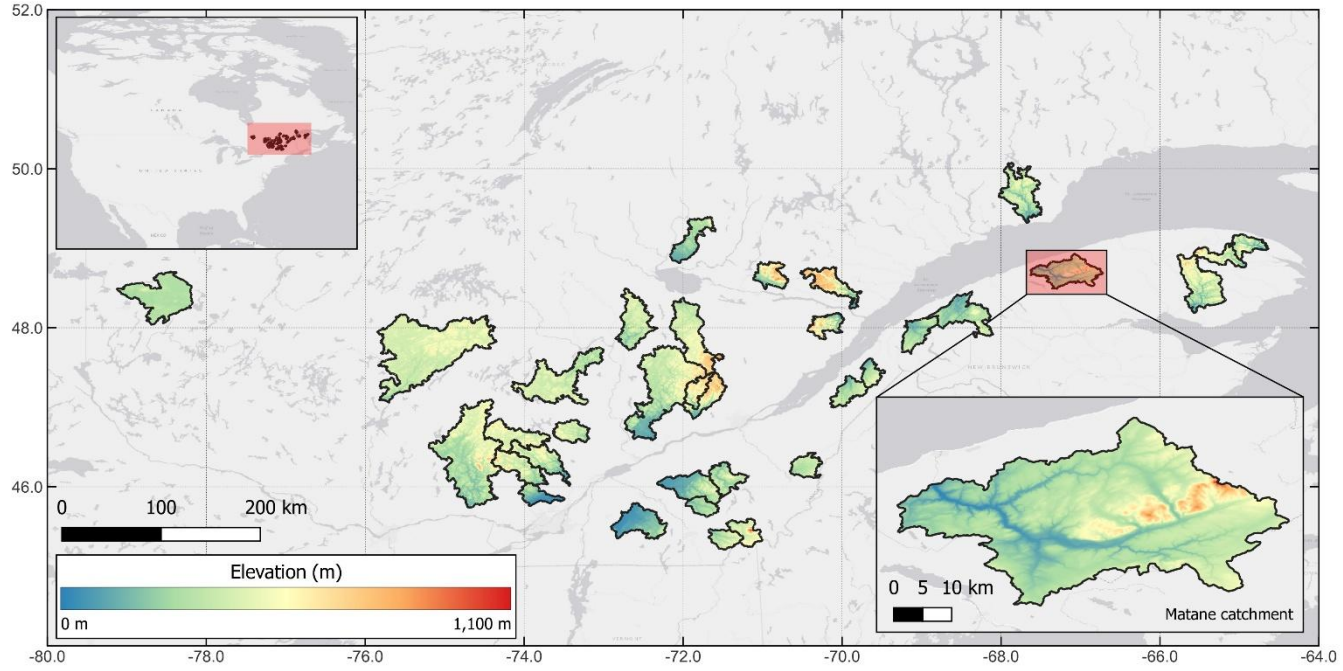


Figure 1. Elevation map of study catchments in southern Quebec.

The Köppen-Geiger Climate Classification designates most of the study area (28 catchments) as belonging to class Dfb (humid continental mild summer, wet all year), except a small part (six catchments) located in the northern portion that belongs to class Dfc (subarctic with cool summers and year-round precipitation) (Beck *et* al., 2018). The region experiences four distinct seasons. Winters are characterized by frequent sub-freezing temperature and significant snowfall. As spring arrives, temperatures gradually rise, leading to significant snowmelt which, along with increasing rainfall, influences streamflow and water availability. Summer brings warmer temperatures, peaking in July, with rainfall remaining relatively high. Fall sees a gradual cooling and a transition from rain to increasing snowfall, setting the stage for another winter cycle. This climatic diversity induces complex hydrological processes at catchment scale, as the interplay between snowmelt and precipitation patterns has a significant influence on streamflow and water availability. These patterns are not unique to Québec but are indicative of broader hydrological changes occurring across boreal regions globally under climate change.

To contextualize the environmental and hydrological setting of the selected catchments, Table 1 presents a synthesis of key descriptors. The table shows the minimum and maximum values for a set of hydrological and geophysical characteristics for each catchment, providing an at-a-glance perspective of the environmental variation within the study area.

Table 1. Hydrological and geophysical characteristics of the study catchments.

Catchment characteristics	Minimum	Maximum
Area (km^2)	525	6840
Mean elevation (m)	137	568
Predominant soil type	Sandy loam	
Predominant land use	Coniferous forest and deciduous forest	
Annual total precipitation (mm)	785	1547
Annual extreme daily temperature (°C)	-37.7	28.6
Annual streamflow ($m^3 s^{-1}$)	10	130

2.2 Data

2.2.1 Hydrometeorological data

This study utilizes meteorological data, specifically total precipitation and mean temperature on a daily time step, sourced from ECMWF's Reanalysis v5 (ERA5) (Hersbach *et* al., 2020). While ERA5 is known to underestimate winter precipitation and exhibit biases in convective precipitation, studies such as Tarek et al. (2020) have demonstrated that ERA5-driven hydrological simulations perform comparably to those using ground-based observational data across Eastern Canada. Their evaluation of 3138 North American catchments found that ERA5-based simulations achieved similar accuracy levels to traditional meteorological observations in hydrological modeling, particularly in Eastern Canada. While observational data can offer higher local accuracy, it also comes with gaps and inconsistencies due to station distribution and measurement errors. ERA5 provided gridded and consistent meteorological inputs across all study catchments, reducing potential biases from heterogeneous station networks. The collected meteorological data spans the period from 1981 to 2020.

Observed streamflow data from 1981 to 2010 was used, recorded at a daily resolution. This data was obtained from the Hydroclimatic Atlas of Southern Québec (MDDELCC, 2022). The dataset contains occasional gaps, primarily during winter months when ice cover and ice jams can significantly distort river flow measurements. To ensure the accuracy of the study, these periods were excluded from the dataset.

2.2.2 Elevation data

A hydrologically conditioned digital surface model was derived from the NASA Shuttle Radar Topography Mission version 3.0 Global 1 (SRTM-DSM) to account for terrain elevation. The SRTM-DSM, originally having a spatial resolution of 30 meters at the equator, underwent resampling to 50 meters resolution and filtering using multiple moving average windows to mitigate the impact of local noise, which could lead to erroneous hydrological behaviours (MacMillan *et* al., 2000). To ensure hydrological consistency, we applied hydrological corrections based on data from provincial agencies (Géobase du réseau hydrographique du Québec (GRHQ) - Données Québec, 2016). To maintain hydrological consistency, we adjusted elevation values along streams by lowering them by 5 meters using the SAGA GIS software (Conrad *et* al., 2015). The resulting DSM accurately captures the hydrological characteristics of the study area and is used for catchment delineation. Additionally, the DSM was resampled to spatial resolutions of 250 and 1000 meters. This resampling process was conducted to optimize computational efficiency while preserving the essential characteristics of the catchments. The minimum value resampling method was used to preserve hydrological connectivity within the study area.

Following this, the Tanalys software (Schulla, 2021) was used to generate key topographic layers, including slope, aspect, and river depth, all formatted for hydrological modeling within WaSiM.

2.2.3 Soil type data

To capture the spatial variability of soil hydraulic properties, we utilized the SIIGSOL 100 meters database (Sylvain et al., 2021), which provides information on soil composition. The SIIGSOL database provides detailed descriptions of the proportions of sand, clay, and silt within the soil profile (MRNF, 2022). In this study, we converted the reported proportions of sand, silt, and clay layers into soil texture classes based on the classification system of the United States Department of Agriculture (USDA). The USDA soil classification system categorizes soils into various texture classes such as loam, clay, sand, silt, and combinations thereof, which are determined based on the percentage composition of each type. This classification aids in understanding the soil's physical characteristics which are crucial factors in hydrological modeling and in predicting soil-water interactions in the studied catchments (Weil and Brady, 2017).

We derived soil hydraulic properties from generated soil type maps, using established relationships between soil texture classes and hydraulic parameters. For the soil type maps, WaSiM generates soil layers of specified thickness based on the control file settings. By default, if there is only one soil type present in the catchment, the soil depth is uniformly distributed throughout the entire area. To account for soil depth variability, we divided soil types into three distinct sections based on their relative elevation within catchment: narrow, normal, and deep. Pixels with elevations below the 33rd percentile were classified as deep, while those with elevations above the 66th percentile were classified as shallow. The remaining soil type rasters fell into the normal category. This classification was based on the imperfect but useful hypothesis that higher elevations correspond to a closer proximity of bedrock to the surface, while lower elevations indicate a greater depth of soil cover in a post-glacial landscape (Akumu *et* al., 2016; Jeong *et* al., 2022).

2.2.4 Land use data

For land use attribution, we used the 2015 North American Land Change Monitoring System (NALCMS) 30 meters land cover dataset (Latifovic *et* al., 2012; Commission for Environmental Cooperation, 2020). The classification scheme used in this map adheres to the widely recognized Land Cover Classification System (LCCS) standard established by the Food and Agriculture Organization (FAO) of the United Nations. This standardized approach ensures the consistency and comparability of land cover information, enabling meaningful regional scale assessments and studies. The nearest neighbor resampling method was employed to align land use maps with the other raster maps used in WaSiM. Land use exerts a substantial influence on various hydrological parameters, and more specifically for the context of this study, it significantly affects parameters such as root distribution, vegetation cover fraction (VCF), roughness length (Z0), and albedo within the hydrological model. The distribution and characteristics of land cover types, ranging from forests to urban areas, directly impact these parameters, thereby influencing processes such as evapotranspiration, runoff, and infiltration.

2.2.5 Groundwater recharge data

In 2008, the Government of Quebec initiated the "Projets d'acquisition de connaissances sur les eaux souterraines" (PACES; roughly translated as "groundwater knowledge acquisition projects") (Carrier *et* al., 2013; Cloutier *et* al., 2013, 2015; Comeau *et* al., 2013; Larocque *et* al., 2013, 2015; Rouleau *et* al., 2013; Buffin-Bélanger *et* al., 2015; Lefebvre *et* al., 2015), aimed at enhancing understanding of the groundwater resources availability in Southern Quebec area. In addition to PACES, numerous studies conducted across the region have estimated groundwater recharge rates, which vary from 50 mm yr^{-1} to over 500 mm yr^{-1} depending on the location and years studied (Croteau *et* al., 2010; Chemingui *et* al., 2015; Larocque *et* al., 2019; Dubois *et* al., 2021; Boumaiza *et* al., 2022).

Of the 34 catchments in this study, fourteen were entirely or partially covered by the PACES project. Table 2 lists these catchments, detailing their areas, associated PACES region reports, the percentage of each catchment's area covered by PACES, and the mean and standard deviation of groundwater recharge for the areas covered.

Table 2. PACES data coverage and groundwater recharge statistics for covered catchments.

Catchment name	Area (km^2)	Region	Cover[1]	PACES recharge	
				Mean (mm yr^{-1})	Std. (mm yr^{-1})
Matane	1650	Bas-Saint-Laurent	31%	179	78
Rimouski	1610	Bas-Saint-Laurent	29%	213	81
Des Trois-Pistoles	932	Bas-Saint-Laurent	38%	74	34
Ouelle	795	Chaudière-Appalaches	62%	180	35
Famine	691	Chaudière-Appalaches	100%	186	46
Bécancour	919	Chaudière-Appalaches and Bécancour	100%	209	83
Nicolet Sud-Ouest	549	Nicolet-Saint-François	100%	242	64
Nicolet	1540	Nicolet-Saint-François	95%	224	82
Noire	1490	Montérégie-Est	93%	133	98
Rouge	5460	Outaouais	26%	310	40
Kinojévis	2590	Abitibi-Témiscamingue	55%	172	87
Petit Saguenay	712	Saguenay-Lac-Saint-Jean	80%	69	78
Petite rivière Péribonca	1090	Saguenay-Lac-Saint-Jean	29%	142	103
Valin	746	Saguenay-Lac-Saint-Jean	73%	221	85
[1] Fraction of total catchment area covered by PACES data.			Median	183	80

2.3 Hydrological modelling

2.3.1 WaSiM model

In this study, we employed WaSiM for hydrological modeling (Schulla, 2021). Hydrological processes were analyzed through three specific configurations: BL (baseline), which serves as the standard comparison model; GW (physical groundwater model), which incorporates detailed groundwater dynamics; and GW-RC (physical groundwater model with constrained recharge), which further refines the groundwater variables by incorporating constrained recharge calibrations. Detailed descriptions of these configurations can be found in Sect. 2.4 of this study.

WaSiM consists of two versions: WaSiM version I, originally developed using the Topmodel approach for simulating subsurface flows based on variable saturation areas, and WaSiM version II, an extended version with the process-oriented Richards approach. The Richards version, which considers hydraulic head gradients and detailed soil physical properties (pF-curve, k(u) function), was selected for this study due to its more physically based nature.

WaSiM follows a modular structure, composed of multiple sub-models that can be activated based on data availability and the specific research objectives. The model operates using a consistent time step, while internally employing flexible sub-

time steps to optimize computational efficiency. It accommodates both regular and irregular raster grids, enabling the analysis of diverse spatial configurations. During each time step, the sub-models are sequentially processed across the entire model grid, enabling parallelization to aid computational optimization and facilitate faster model execution.

One of the key process modules within WaSiM is the unsaturated zone model, which plays a crucial role in calculating various hydrological variables such as surface runoff, groundwater recharge, interflow, and baseflow. Interflow refers to water moving laterally through the upper soil layers, contributing to streamflow, while baseflow is the portion of streamflow sustained by groundwater flow. These variables are essential for understanding the water balance and hydrological dynamics within the study area. Table 3 provides an overview of the hydrological model configuration used in this study.

Table 3. Overview of WaSiM characteristics and sub-models used in this study.

Sub-model	Method	Reference
Meteorological interpolation	Inverse distance interpolation	(Shepard, 1968)
Potential evapotranspiration	Hamon approach	(Hamon, 1963)
Actual evapotranspiration	Richards equation using the Van Genuchten parameters	(Richards, 1931; van Genuchten, 1980)
Snow melt	Temperature-index approach	(Hock, 2003)
Interception	Classic bucket approach dependent on LAI	-
Lake modelling	Integrated approach to model natural and artificial lakes, considering interactions with unsaturated zone, routing, snow, evaporation, interception, and groundwater models.	-
Unsaturated zone flow	Richards equation using the Van Genuchten parameters	(Richards, 1931; van Genuchten, 1980)
Groundwater flow	Integrated two-dimensional groundwater model	-
Routing	Kinematic wave approach	(Lighthill and Whitham, 1955)

Meteorological data interpolation was an essential step in the hydrological modeling process. The chosen hydrological model, WaSiM, performed the interpolation of daily precipitation and temperature inputs between ERA5 points. For each simulation, the model creates grids that incorporate the interpolated meteorological values at the model's spatial resolution, effectively representing the climatic conditions for each individual pixel. The inverse distance weighting method was used as recommended by WaSiM model description report (Schulla, 2021).

2.3.2 Calibration parameters

Calibration of WaSiM involved the optimization of 17 parameters, selected in accordance with WaSiM documentation (Schulla, 2021), while the remaining parameters in the control file were set to their default values. Table 4 provides a

235 detailed description of upper and lower limits set for calibrating the 17 parameters in WaSiM, with each parameter adjusted to two decimal places within the specified calibration range.

Table 4. Description of the parameters used for the calibration of WaSiM.

No.	Code	Description	Unit	Sub-Model	Range
1	k_D	Storage coefficient for surface runoff	h	Unsaturated zone	[1, 25]
2	k_H	Storage coefficient for interflow	h	Unsaturated zone	[1, 25]
3	d_r	Drainage density for interflow	m^{-1}	Unsaturated zone	[1, 50]
4	QD_{Snow}	Fraction of surface runoff on snow melt	-	Unsaturated zone	[0.1, 1]
5	c_0	Degree-Day factor	$mm°C^{-1}\ d^{-1}$	Snow	[0, 3]
6	T_0	Temperature limit for snow melt	°C	Snow	[-4, 4]
7	$T_{R/S}$	Transition temperature snow/rain	°C	Snow	[-4, 4]
8	C_{WH}	Water storage capacity of snow	-	Snow	[0.1, 0.3]
9	C_{rfr}	Coefficient for refreezing	-	Snow	[0.1, 1]
10	$f_{i,summer}$	Summer correction factors for ETP	-	Evapotranspiration	[0.1, 2]
11	$f_{i,fall}$	Fall correction factors for ETP	-	Evapotranspiration	[0.1, 2]
12	$f_{i,winter}$	Winter correction factors for ETP	-	Evapotranspiration	[0.1, 2]
13	$f_{i,spring}$	Spring correction factors for ETP	-	Evapotranspiration	[0.1, 2]
14	K_{rec}	Recession constant for hydraulic conductivity	-	Soil table	[0.1, 0.99]
15	d_z[a]	Soil layer thickness	-	Soil table	[0.8, 1.4]
16_A	K_B	Storage coefficient for base flow	m	Unsaturated zone	[0.1, 8]
17_A	Q_0	Scaling factor for base flow	$mm\ h^{-1}$	Unsaturated zone	[0.1, 5]
16_B	Kol[b]	Colmation of the river links	-	Input grid	[1, 100]
17_B	K_{XY}[c]	Saturated horizontal conductivity (x-y-direction)	$m\ s^{-1}$	Input grid	[0.2, 4]

[a] Calibration coefficient, ranging from 0.8 to 1.4, is applied to adjust the total soil depth, which is predetermined to be 8 meters for shallow, 14 meters for normal, and 20 meters for deep soil conditions.

[b] Calibration coefficient, ranging from 0.8 to 1.4, is applied to adjust the colmation grid, which is predetermined to be $1x10^{-6}$.

[c] A calibration coefficient, ranging from 0.2 to 4, is applied to adjust the saturated horizontal conductivity grid, which is predetermined to be $4x10^{-5}$ $m\ s^{-1}$.

For each model configuration (BL, GW, GW-RC), the full set of 17 parameters was recalibrated independently within the
240 specified ranges. In the BL configuration, where the groundwater model is not activated and groundwater flow is computed using a conceptual approach within the unsaturated zone sub-model, parameters 16_A (K_B) and 17_A (Q_0) were the ones calibrated for baseflow representation, as these parameters are only relevant when the conceptual groundwater scheme is used. Groundwater flow is assessed using Eq. (1) (Schulla, 2021), which calculates baseflow as a function of several parameters including the scaling factor for baseflow (Q_0) and the recession constant for baseflow (K_B).

$$Q_B = Q_0 * K_s * e^{(h_{GW} - h_{geo,0})/K_B}, \quad (1)$$

where Q_B is baseflow (m s^{-1}), Q_0 is a scaling factor for baseflow, K_s is the saturated hydraulic conductivity (m s^{-1}), h_{GW} is the groundwater table height (m), $h_{geo,0}$ is the geodetic altitude of the soil surface (m) and K_B is the recession constant for baseflow (m).

In the configurations used in GW and GW-RC, which activate groundwater model, parameters 16_A and 17_A are replaced by parameters 16_B and 17_B to obtain a more physically based representation of groundwater processes. Parameters 16_B and 17_B adjust values associated to two input grids that allow to account for the colmation of the river links and saturated horizontal conductivity. This distinction ensures a consistent number of calibrated parameters across all configurations, facilitating an unbiased comparison of model performance.

2.3.3 Model optimization

Parameters optimization was performed independently for each catchment through the dynamically dimensioned search algorithm (DDS; (Tolson and Shoemaker, 2007)), following the recommendation of Arsenault *et* al. (2014). This algorithm is specifically designed for efficiently calibrating complex hydrological models with a large parameter range given a finite computing budget. During optimization, it dynamically adapts its search strategy based on the number of evaluations performed and performance metrics. To manage computational demands effectively while ensuring thorough exploration of the parameter space, a two-phase calibration strategy was employed, albeit the approaches differ for the constrained groundwater configurations.

Initially, 1000 simulations were performed for each catchment at a broader spatial resolution (1000 meters) using a broader range of values for each parameter (Table 4). This phase aimed to identify an approximation of the optimal values for each parameter. Subsequently, these values were used to initialize the second calibration step at a finer spatial resolution (250 meters). This two-step approach was chosen based on preliminary testing on the Bonaventure and Matane catchments, which demonstrated that transferring optimized parameters from 1000 m resolution to 250 m required only minor refinements. Additional tests showed that increasing the number of simulations at 250 m resolution beyond 50 runs (e.g., 75 or 100) provided negligible improvements in model performance, making further computational expense unjustified. This sequential calibration strategy allows to refine the model's performance progressively. By first identifying a set of parameters that achieves reasonable model performance at a coarser scale, we then fine-tune the model at a higher resolution to enhance the spatial distribution of hydrological simulations.

The objective functions used vary by configuration: For BL and GW, the objective is to optimize the Kling-Gupta Efficiency (KGE, (Kling *et* al., 2012)), as discussed in Sect. 2.5.1. Conversely, the GW-RC configuration employs a modified objective function that seeks to optimize KGE and constrain groundwater recharge rates and variability. This approach is described in Sect. 2.4.3 and Sect. 2.5.2.

The study employed split-sample test (SST) framework for the parameter optimization assessment. This widely used approach involves dividing the available data into two sets: one for calibrating the model and the other for validating its

performance on unseen time periods. The calibration period (2000-2009) and the validation period (1990-1999) were chosen based on the availability of comprehensive and reliable hydrological data. To minimize the impact of missing streamflow data, calibration and validation years were selected to ensure that most catchments had complete records. However, data gaps were noted for three catchments: Croche, Petit Saguenay, and Sainte-Marguerite Nord-Est. Specifically, Croche lacked data from 2001 to 2004, Petit Saguenay from 2000 to 2010, and Sainte-Marguerite Nord-Est from 1998 to 2010. To accommodate these gaps, adjustments were made to the calibration and validation periods for the affected catchments. The calibration periods were shortened to later years: 1995 to 1999 for Croche and Petit Saguenay, and 1992 to 1996 for Sainte-Marguerite Nord-Est. Correspondingly, the validation periods were adjusted to precede the missing data: 1991 to 1994 for Croche, 1986 to 1994 for Petit Saguenay, and 1986 to 1991 for Sainte-Marguerite Nord-Est. A five-year spin-up period was performed before each simulation to allow the model to reach a stable state, eliminating the influence of unstable initial conditions on the model's performance metrics.

2.4 Model configurations

The primary objective of this research is to examine how different model configurations influence the representation of hydrological processes. To ensure a consistent comparison of model configuration and calibration, we designed a modelling framework that allow to compare three configurations that incrementally incorporate more complex hydrological variables.

2.4.1 Baseline

The first configuration (BL), serving as baseline configuration, employs the standard calibration of the model without activating the groundwater module. This configuration is aligned with the traditional application of WaSiM, where the focus is predominantly on streamflow, and groundwater flow is modeled using Eq. (1) within the unsaturated zone sub-model. This configuration is comparable to what has been frequently adopted in numerous studies, providing a common basis for comparative analysis (Rössler *et* al., 2012; Förster *et* al., 2018; Markhali *et* al., 2022; Valencia Giraldo *et* al., 2023).

2.4.2 Physical groundwater module

The second configuration, GW (physical groundwater), marks a departure from the BL configuration by activating WaSiM's groundwater module. This adjustment allows for groundwater flow to be simulated within a designated sub-model, transitioning from a conceptual to a more physically based representation. In WaSiM, the groundwater model is coupled bi-directionally with the unsaturated zone, ensuring a dynamic exchange of water fluxes. The unsaturated zone module calculates fluxes between the unsaturated zone and the groundwater that act as the upper boundary condition for the groundwater model, while the groundwater module simulates lateral flow and adjusts the groundwater table, feeding back changes to the unsaturated zone as inflow or outflow. This configuration, used in numerous studies (Bormann and Elfert, 2010; Natkhin *et* al., 2012; Gädeke *et* al., 2014; Schäfer *et* al., 2023), is recommended by the WaSiM documentation for

catchments where groundwater dynamics play a pivotal role in the hydrological cycle, particularly in lowland areas with extensive sediment layers.

2.4.3 Physical groundwater module and constrained recharge

For configuration GW-RC (physical groundwater and constrained recharge), we incorporate groundwater recharge into the calibration process to achieve a better representation of hydrological variables such as baseflow, interflow, and runoff. Importantly, GW-RC uses the same model structure as GW, with the goal of isolating the effect of adding groundwater recharge in calibration. By introducing recharge into the calibration, we restrict hyperplane exploration and ensure that the model's representation of the hydrological cycle is more accurately simulating groundwater recharge dynamics. This is particularly useful if model hydrological variables are an important input to another analysis or process, such as for better understanding groundwater movement and evolution under climate change for certain types of vegetation, for example.

GW-RC calibration was performed in two phases. First, we defined new parameter ranges for variables affecting baseflow (dr, QDSnow, Krec, Kol, Kxy). We first conducted 200 evaluations at a spatial resolution of 1000 meters, followed by 50 evaluations at 250 meters using the objective function presented in Eq. (2). Essentially, the aim here is to constrain the parameter set to a single value that performs well overall and provides realistic internal variables. Similar approaches have been used in studies such as Duethmann *et* al. (2024), which underscores the benefits of integrating Landsat-derived land surface temperature (Ts) data into model calibration. Landsat, a series of Earth-observing satellites, provides crucial Ts data used in this study. By including satellite-derived Ts, the study demonstrated improvements in the model's ability to capture spatial anomalies and ecosystem stress responses, while maintaining streamflow accuracy, illustrating the advantages of multi-variable constraints in model calibration.

Following pre-calibration at both spatial resolutions, the resulting calibrated parameter sets were analyzed to define new parameter ranges for the calibration phase. This analysis involved adjusting the minimum and maximum values of parameters influencing baseflow (d_r, QD_{Snow}, K_{rec}, Kol, Kxy) by ±10% to establish new calibration ranges.

In the second and most important calibration phase, the process continued with the adjusted parameter ranges, employing a less restrictive objective function (Eq. (3)) to better accommodate uncertainties in the recharge data. This phase involved a comprehensive series of 1000 evaluations at 1000 meters and 50 at 250 meters resolutions. The modified objective function primarily emphasized the KGE while incorporating the standard deviation of recharge at a reduced influence of 4%. This modification was crucial to allow the model flexibility to adapt the groundwater recharge rate according to the specific hydrological characteristics and precipitation patterns of each catchment. Given that the initial recharge rate of 250 mm yr^{-1} was a preliminary estimate and not necessarily reflective of individual catchment conditions, this approach enabled a more tailored calibration.

A key justification for not applying the same constrained parameter range across all configurations is that BL and GW do not incorporate recharge in calibration. Their parameters optimization is based solely on streamflow, whereas GW-RC explicitly integrates recharge to constrain the parameters range.

Table 5 shows an overview of the three methods to ease comparisons between configurations.

Table 5. Summary of configurations

Settings	BL	GW	GW-RC
Groundwater Modelling	Conceptual within unsaturated zone sub-model	Physically based within the groundwater sub-model	Physically based within the groundwater sub-model
Calibration Parameters	17 parameters (including K_B and Q_0)	17 parameters (including Kol and K_{XY})	17 parameters (including Kol and K_{XY})
Precalibration	N/A	N/A	200 simulations at 1000 meters followed by 50 simulations at 250 meters
Calibration	1000 simulations at 1000 meters followed by 50 simulations at 250 meters	1000 simulations at 1000 meters followed by 50 simulations at 250 meters	1000 simulations at 1000 meters followed by 50 simulations at 250 meters
Objective function	Kling-Gupta efficiency	Kling-Gupta efficiency	Constrained Kling-Gupta efficiency
Computational demand	10 CPU-year at 4.5 GHz	10 CPU-year at 4.5 GHz	15 CPU-year at 4.5 GHz

CPU-year : A CPU-year is the effort of a CPU running for one year.

2.5 Performance assessments

The KGE (Kling *et* al., 2012) was chosen as the objective function to assess the model's performance during the calibration process of configurations BL and GW.

An arbitrary baseline groundwater recharge rate of 250 mm yr^{-1} and a standard deviation of 80 mm yr^{-1} have been established as representative benchmarks for the studied catchments. These values are based on PACES data and additional studies conducted in Quebec, as described in Sect. 2.2.5. The objective function for the pre-calibration of configuration GW-RC, outlined in Eq. (2), aims to balance KGE with these established recharge metrics. Specifically, the function assigns a weight of 70% to KGE, 20% to the annual recharge standard deviation, and 10% to the mean annual recharge. This specific weighting was determined based on preliminary testing conducted on two test catchments, where various weight combinations were evaluated. The selected weights provided the best trade-off, ensuring that recharge estimates remained realistic while maintaining strong KGE values for streamflow. In particular, assigning 20% to the recharge standard deviation and 10% to the mean annual recharge allowed the model to better capture recharge variability without compromising overall streamflow performance. This objective function was designed to ensure both the quantity and variability of recharge were realistically modeled without sacrificing performance in terms of overall streamflow quality through the KGE.

The objective function employed in the pre-calibration of GW-RC configuration is formulated as follows:

$$Precalibration\ function = 1 - (0.7 * KGE + 0.2 * \left|\sigma_{r_{sim}} - 0.08\right| + 0.1 * \left|\overline{r_{sim}} - 0.25\right|), \quad (2)$$

where $\sigma_{r_{sim}}$ is the simulated annual recharge standard deviation (m yr^{-1}), $\overline{r_{sim}}$ is the simulated mean annual recharge (m yr^{-1}) and KGE is the Kling-Gupta efficiency.

Groundwater recharge simulations were performed at the pixel level, ensuring detailed local representation. The simulated mean annual recharge reflects the average amount of recharge occurring annually across the entire catchment during the calibration period. Similarly, the simulated annual standard deviation quantifies the variability in annual recharge across all pixels within the catchment during the same period. Introducing pixel level standard deviation helps in curbing extreme values in groundwater recharge, thus stabilizing the simulation outputs. The mean annual recharge is employed to verify that the model accurately captures the overall recharge volume expected for the study area.

For the main calibration phase of the GW-RC configuration, the objective function is simplified to focus more intensively on streamflow accuracy:

$$Calibration\ function = 1 - (0.96 * KGE + 0.04 * \lceil \sigma_{r_{sim}} - 0.08 \rceil), \quad (3)$$

where $\sigma_{r_{sim}}$ is the annual recharge standard deviation (m yr^{-1}) and KGE is the Kling-Gupta efficiency.

2.6 Statistical analysis

To assess the performance of the hydrological model configurations, statistical analyses were conducted to compare calibration and validation performance across different configurations. The primary metric used was the KGE, which evaluates the accuracy of simulated streamflow against observed data. The performance metrics were analyzed for each configuration during both the calibration period (2000-2009) and validation period (1990-1999), ensuring robust evaluation across varying hydrological conditions.

All statistical comparisons were made using the Kruskal-Wallis test, a non-parametric method chosen due to its suitability for non-normally distributed data. This test was employed to detect significant differences in the performance and hydrological responses between the model configurations. Where significant differences were identified, multiple comparison post-hoc tests were conducted to ascertain the specific pairs of configurations that differed significantly.

Pearson's correlation coefficients were used to explore the influence of calibration parameters on hydrological variables. This statistical approach provided insights into how variations in parameter settings across different configurations could affect the representation of hydrological processes like surface runoff, interflow, and groundwater recharge.

3 Results

3.1 Calibration and validation performance

Throughout the calibration (2000-2009) and validation (1990-1999) periods, all configurations yielded KGE values above 0.5. Calibration and validation performances were very similar, with a deviation less than 5%, demonstrating the robustness

of the simulations. KGE values for all catchments and configurations, for both the calibration and validation periods, are presented in Table A1.

Figure 2 reveals a clear trend where catchments with high KGE values during calibration tend to maintain similar performance during validation. This consistency underpins the robustness of the configurations across different validation periods. During the validation period, median KGE values were higher for configurations BL (0.824) and GW (0.830) compared to GW-RC (0.770), demonstrating superior performance in the models without groundwater recharge constraints. However, GW-RC demonstrates more consistent KGE values between calibration and validation, suggesting it may offer more stability in model performance despite its slightly lower KGE scores.

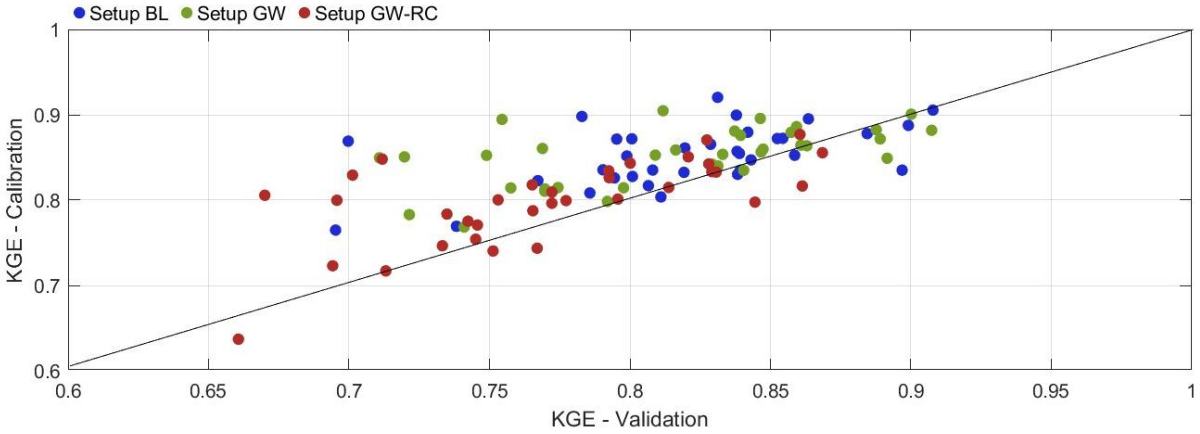


Figure 2. Comparison of Kling-Gupta Efficiency values between calibration and validation periods for three configurations. Each point represents a catchment, color-coded by configuration: Configuration BL (blue), Configuration GW (green), and Configuration GW-RC (red). The line represents a one-to-one relationship where calibration and validation KGE values are equal. Points below the line indicate better performance in the validation phase compared to calibration, while those above the line show a decline in performance from calibration to validation.

It is important to note that the KGE values for configuration GW-RC are slightly lower than those from configurations BL and GW, which is expected given the supplementary constraints imposed during calibration.

3.2 Hydrological variables analysis

This section delves into the simulated hydrological variables, examining their range and distribution across the various model configurations during the calibration and validation periods. The variables in focus include surface runoff, baseflow, interflow, groundwater recharge, and actual evapotranspiration (ETa).

Figure 3 illustrates the annual totals (means for groundwater level and soil moisture) for simulated hydrological variables for both calibration and validation periods and for all catchments. Notably, there is a consistency in the distribution of hydrological variables of each model configuration between the calibration and validation periods. This allows us to focus our detailed analysis solely on the validation period for conciseness.

A comparative assessment reveals distinct patterns in the simulated hydrological variables among the configurations. Specifically, configuration GW-RC simulates higher surface runoff and lower interflow, and infiltration compared to configurations BL and GW. Conversely, configuration BL is characterized by higher actual evapotranspiration, lower groundwater recharge, and a higher groundwater level. Configuration GW shares similarities with both configuration BL (in terms of runoff, interflow, and infiltration) and configuration GW-RC (regarding baseflow, groundwater recharge, actual evapotranspiration, and groundwater level).

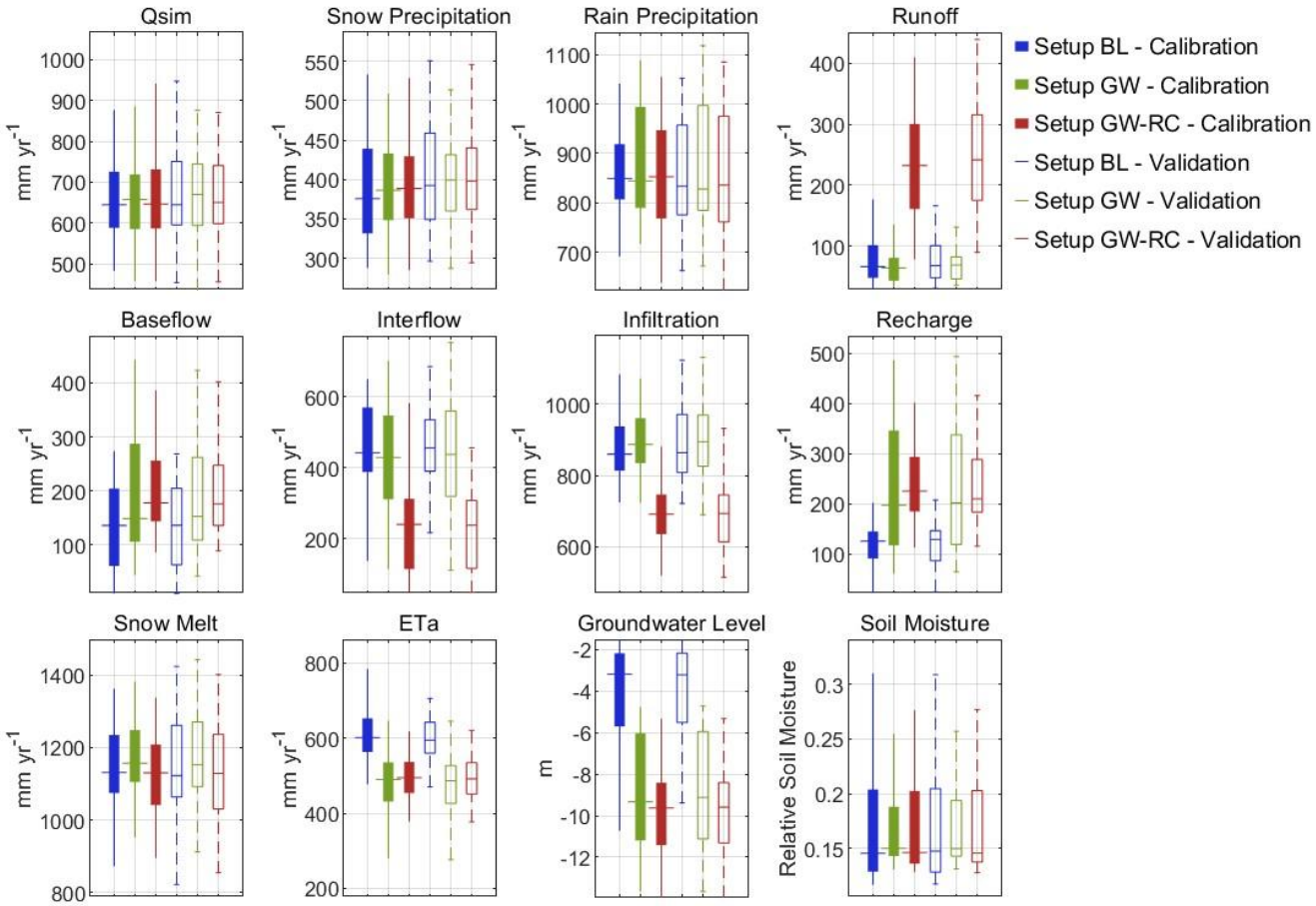


Figure 3. Boxplots illustrating annual totals (means for groundwater level and soil moisture) variability of model internal variables. These boxplots detail the variability of key hydrological variables modeled with the different configurations, for calibration and validation periods and for all catchments.

Figure 4 presents the proportional distribution of surface runoff, baseflow, interflow, and actual evapotranspiration for the three hydrological model configurations (BL, GW, and GW-RC). The charts effectively compare the relative contribution of each process to the total water cycle within the modeled catchments.

The figure highlights that configuration GW-RC simulates a notably higher proportion of surface runoff (21%) and baseflow (17%) with a lower proportion of interflow (20%). Conversely, configuration BL has a higher proportion of actual evapotranspiration (47%) and less baseflow (11%). Finally, configuration GW has similarities with both BL (surface runoff

and interflow) and GW-RC (baseflow and actual evapotranspiration) configurations. The factors influencing the differences between configurations are further analyzed in the discussion section.

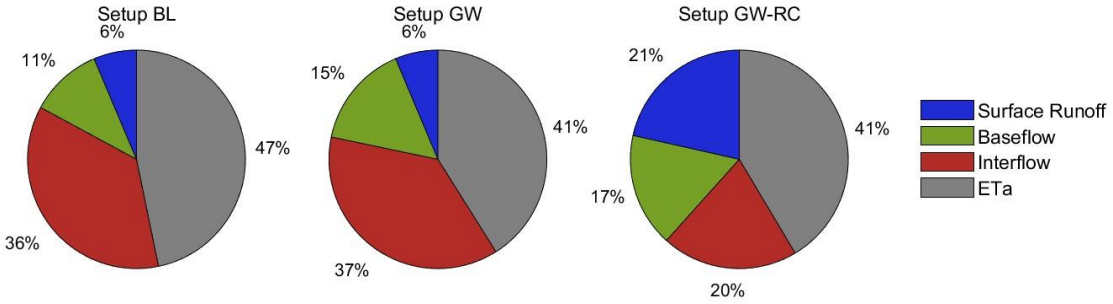


Figure 4. Proportional distributions of key hydrological variables for the BL, GW and GW-RC hydrological model configurations for the validation period (1990-1999).

Table 6 shows that the observed similarities in surface runoff and interflow between configurations BL and GW are substantiated by statistical significance in their mean groupings. Furthermore, the parallels drawn between configurations GW and GW-RC in terms of actual evapotranspiration and groundwater recharge are also supported by significant statistical evidence. However, the apparent similarity in baseflow between configurations GW and GW-RC does not hold statistical significance. This outcome is expected, as both GW and GW-RC employ the same groundwater module, with GW-RC differing only in its calibration approach. The observed variations in baseflow arise from the inclusion of recharge constraints in GW-RC. More broadly, the significant contrast in baseflow between BL and the other two configurations suggests that the choice of model configuration plays a primary role in determining baseflow dynamics rather than the specific calibration strategy applied.

Table 6. Statistical analysis of the differences in estimated hydrological variables from the three configurations BL, GW and GW-RC.

Hydrological Variables	BL vs. GW	BL vs. GW-RC	GW vs. GW-RC
Surface runoff	0	1	1
Baseflow	1	1	1
Interflow	0	1	1
Actual evapotranspiration	1	1	0
Groundwater recharge	1	1	0

(Not Different = 0; Different = 1)

Figure 5 illustrates the annual totals distribution of key hydrological variables (surface runoff, baseflow, interflow, actual evapotranspiration, groundwater recharge, and precipitation) across 34 catchments for each model configuration (BL, GW, and GW-RC). The figure provides a comprehensive comparison of how each configuration partitions the water balance components for each catchment. Consistent trends in hydrological responses are observed across the catchments for each model configuration. For instance, configuration GW-RC shows higher surface runoff and baseflow, with lower interflow values compared to the other configurations indicating that calibration strategies and model complexity influence the distribution of water fluxes. In contrast, configuration BL consistently reports higher actual evapotranspiration (ETa) and lower groundwater recharge. Statistical comparisons indicated that baseflow, surface runoff and interflow dynamics of GW-RC configuration are significantly different compared to BL and GW configurations (Table 6).

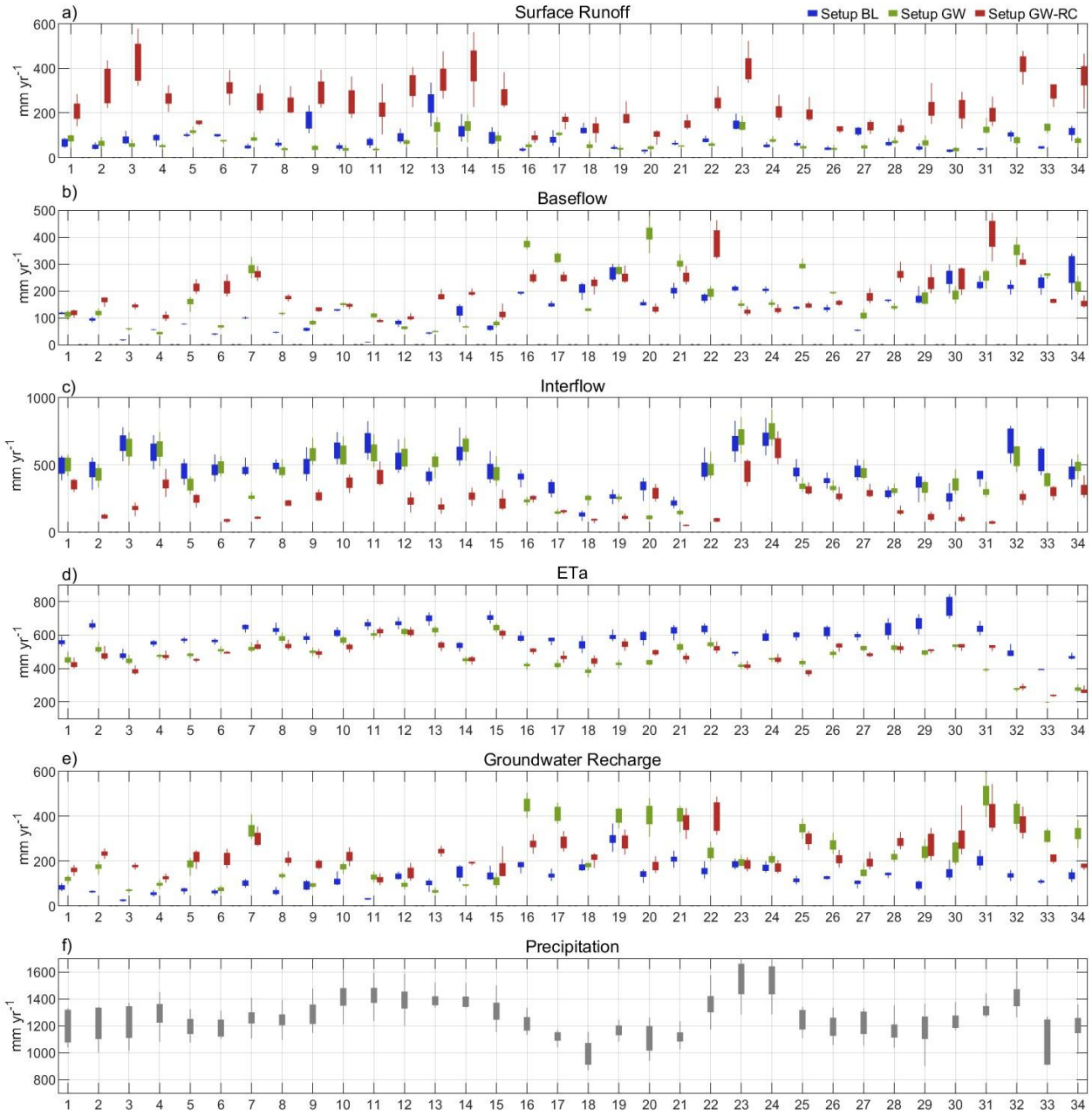


Figure 5. Boxplots of annual values for key hydrological variables predicted by WaSiM for the 34 catchments and three configurations for the validation period (1990-1999).

3.3 In-depth analysis of the Matane catchment

This section explores the temporal dynamics of streamflow and hydrological variables in the Matane catchment, which was selected as a representative example from the study's catchments. Figure 6 reveals consistent patterns in hydrological variable behavior across all configurations during both the calibration and validation periods. Consequently, the following discussions will focus primarily on the validation period. Generally, interflow is the major contributor to simulated streamflow in configurations BL and GW throughout the year. In contrast, configuration GW-RC is characterized by a significant increase in surface runoff during the seasonal high flow and high precipitation periods in the fall, while predominantly exhibiting interflow contributions during other times of the year.

Configuration GW-RC is also marked by higher levels of surface runoff and baseflow, but lower interflow compared to the other configurations. Configuration BL is distinguished by having the highest levels of annual actual evapotranspiration. Configuration GW aligns closely with configuration BL in terms of interflow, surface runoff, and baseflow, demonstrating similar hydrological dynamics between these two configurations.

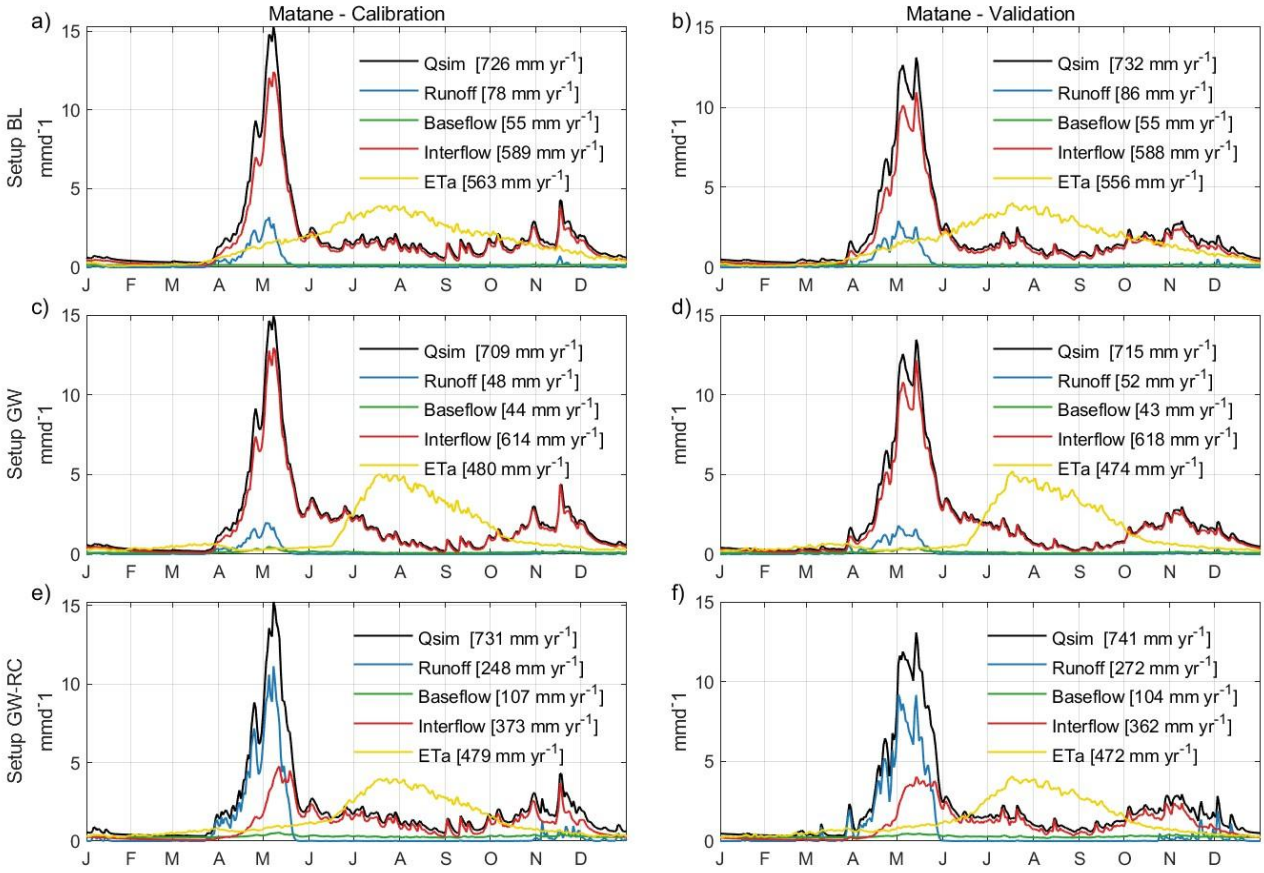


Figure 6. Detailed hydrological variable hydrograph for Matane catchment during both the calibration and validation phases and for the three configurations. Calibration results are shown in panels (a), (c), and (e) for Configurations BL, GW, and GW-RC, respectively, while validation results are depicted in panels (b), (d), and (f). These hydrographs demonstrate how baseflow, interflow and runoff contribute to total streamflow throughout the year, with noted annual totals provided for a comprehensive comparison.

Figure 7 reveals seasonal variations that correlate with hydrological responses to climatic conditions. Surface runoff and interflow differ significantly during periods of high flow, typically driven by snowmelt. Configurations BL and GW primarily attribute high flows to interflow, whereas configuration GW-RC reflects these peaks with increased surface runoff. Groundwater recharge in configuration BL exhibits more pronounced seasonal fluctuations compared to the patterns observed in configurations GW and GW-RC. Similarly, configuration BL maintains a consistent baseflow year-round, unlike configurations GW and GW-RC, which show seasonal baseflow variations. In terms of actual evapotranspiration,

configuration BL consistently exhibits higher rates in the spring and fall, GW peaks during the summer, and GW-RC displays a pattern that blends characteristics of both BL and GW across different seasons.

Figure 7 panel C illustrates the daily groundwater recharge in the Matane catchment for each configuration. A common seasonal pattern is evident across all configurations: recharge decreases in winter, rises significantly during snowmelt, and then exhibits marked variability throughout summer and autumn. Notably, configuration GW-RC shows a lower dynamic range during snowmelt compared to configurations BL and GW, which exhibit more pronounced peaks. Throughout the winter, summer, and autumn months, configuration GW-RC consistently shows higher recharge rates than the other configurations. The trends observed in the Matane catchment are also representative of the behaviors seen across all studied catchments.

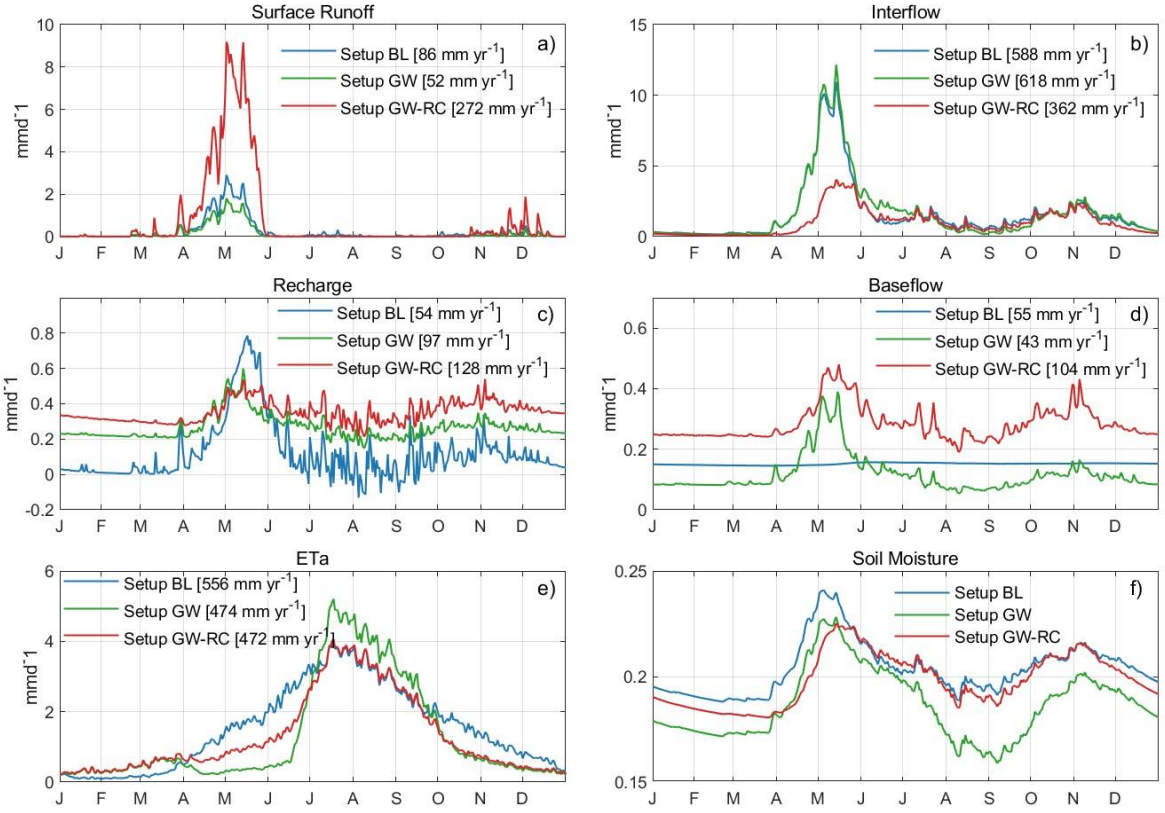


Figure 7. Seasonal distribution of hydrological variables in the Matane catchment for the validation period (1990-1999). This figure visualizes the annual distribution of key hydrological variables across the three configurations throughout the year.

4 Discussion

4.1 Performance against representation

This study aimed to analyze how varying model configurations affect the representation of hydrological variables estimated by WaSiM. Through the comparative analysis of three distinct calibration configurations, BL (baseline model), GW (activated groundwater simulation), and GW-RC (groundwater simulation and recharge calibration), this study provides insights into how internal hydrological processes are represented in a physically based model.

KGE values were consistently higher for the BL and GW configurations compared to GW-RC during both calibration and validation periods. Configuration GW-RC's modestly lower performance on KGE is reflective of its calibration not solely focusing on optimizing KGE but also in incorporating a broader suite of hydrological dynamics.

This finding aligns with prior research, which suggests that adding constraints to model parameters can often improve the representation of other hydrological processes, such as groundwater dynamics and soil moisture, albeit at the cost of lower validation performance. For instance, Yassin *et* al. (2017) emphasized that incorporating additional data, such as from the Gravity Recovery and Climate Experiment (GRACE), can lead to more comprehensive and physically realistic model. Similarly, Dembélé *et* al. (2020) showed that incorporating spatial patterns from satellite data significantly improve the model's representation of soil moisture and evapotranspiration. Similarly, Bouaziz *et* al. (2021) found substantial disparities in internal process representation among models calibrated to the same streamflow data, highlighting the limitations of relying solely on discharge data for model validation. Lastly, Pool *et* al. (2024) demonstrated that incorporating variables such as actual evapotranspiration and total water storage alongside discharge in model calibration can significantly enhance the simulation accuracy for these variables.

4.2 Hydrological variables analysis

Regarding the distribution of hydrological variables, configuration BL demonstrated the highest actual evapotranspiration rates, alongside the lowest groundwater recharge and baseflow. Conversely, GW-RC was noted for the highest surface runoff and the lowest interflow. Configuration GW exhibited characteristics that were intermediate between the other two configurations. It resembled BL in terms of interflow and surface runoff but aligned more closely with GW-RC for groundwater recharge, actual evapotranspiration, and baseflow.

As shown in Figure C1, baseflow is closely correlated ($r = -0.875$) with the drainage density parameter (scaling parameter for interflow) for configurations GW and GW-RC. The constrained parameter range in configuration GW-RC explains the minor differences in baseflow rates observed between these configurations. In contrast, the baseflow in configuration BL is significantly correlated ($r = 0.715$) with the scaling factor for baseflow. The differences in groundwater recharge and baseflow across the configurations can be primarily attributed to the activation of the groundwater flow sub-model. In WaSiM, the simulation of groundwater processes can either follow a more conceptual or physically based pathway. Our

results indicated that GW and GW-RC, which incorporate more complex mechanisms between groundwater and surface processes, lead to more dynamic and possibly more accurate representations of baseflow and recharge dynamics.

The disparities in interflow for configuration GW-RC are primarily due to the restricted calibration of the drainage density parameter. A strong correlation (r = 0.801) between interflow rates and the parameter value highlights how constraining groundwater recharge during calibration can influence other hydrological variables, such as interflow. Variations in surface runoff for configuration GW-RC are linked to calibration restrictions on the 'QDsnow' parameter, which represents the fraction of surface runoff from snowmelt. A strong correlation (r = 0.899) between this parameter and surface runoff rates indicates that it has a significant influence on this hydrological variable. Also, configuration GW-RC showed the highest value for 'QDsnow' parameter and the lowest value for the drainage density parameter consequently leading to the highest surface runoff and lowest interflow rates. This observation indicates that interflow is a flexible variable within the model, with configurations BL and GW appearing to prioritize it over surface runoff and baseflow. This prioritization allows the optimization algorithm greater latitude to enhance performance metrics like KGE and more accurately reproduce observed streamflow patterns. Conversely, configuration GW-RC, constrained by groundwater recharge, tends to prioritize baseflow and surface runoff. While this approach may reduce the model's flexibility in mirroring observed streamflow, it enhances the precision with which other hydrological processes are represented as detailed in Sect. 4.3. The same trend was found for the Matane catchment, underlining the broader applicability of these findings across different geographical contexts. Such a representation offers essential information that can be pivotal for water management strategies.

Moreover, configuration GW-RC also exhibited lower values of k_h (storage coefficient for interflow), higher values of Krec (recession constant for hydraulic conductivity), lower correction factors for PET in summer, and higher correction factors for PET in winter compared to the other two configurations. These differences indicate that adding groundwater recharge constraints during calibration can influence parameter values in sub-models that are seemingly unrelated to groundwater processes, such as evapotranspiration. This suggests that the recharge constraint propagates through the model structure, affecting multiple hydrological components. A complete list of calibrated parameter values for each catchment and configuration is provided in Appendix D.

4.3 Pinpointing the optimal model configuration

The differences in surface runoff during the snowmelt season across configurations can be largely attributed to the parameter QDsnow. WaSiM employs a singular parameter (QDsnow) to account for surface runoff from snowmelt. This parameter is calibrated between 0 and 1, and its precise setting critically influences the model's surface runoff predictions.

Analysis of Fig. 6 reveals that configurations BL and GW exhibit lower surface runoff from snowmelt, where melted snow predominantly percolates into the soil, contributing to interflow rather than surface runoff. This behavior is unexpected because, in fully frozen soil conditions, significant surface runoff is typically anticipated due to reduced infiltration.

Conversely, configuration GW-RC, which integrates groundwater recharge into the calibration process, follows a more typical hydrological pattern. Higher surface runoff is observed at the onset of snowmelt, gradually decreasing as infiltration

and interflow increase when the soil thaws. This progression aligns with the expected hydrological responses in frozen terrains, illustrating how the inclusion of groundwater recharge can improve the model's simulation of seasonal transitions. This trend of higher surface runoff during snowmelt was observed consistently across all catchments in the study, with detailed figures provided in the supplementary material (Fig. S1 to Fig. S32). Configuration GW-RC showed increased surface runoff during the snowmelt period compared to the other configurations. However, for 11 out of the 34 catchments, the surface runoff results were notably elevated. Figure B1 illustrates an example where nearly all of the spring discharge was attributed to surface runoff, suggesting that the value assigned to the QDsnow parameter, when set too close to 1, may lead to an overestimation of runoff. Careful calibration of this parameter is essential to avoid misrepresentations in the hydrological processes.

The analysis of groundwater recharge, as detailed in Sect. 3.4, reveals significant differences in seasonal dynamics and spatial distribution among the configurations. Notably, GW-RC displays less dynamic recharge rates during the snowmelt period compared to configurations BL and GW. This is indicative of a distinct interplay between surface runoff and infiltration processes within configuration GW-RC, where higher surface runoff during the spring results in reduced infiltration. Additionally, GW-RC exhibits higher recharge rates during summer, fall, and winter, with a peak in fall.

The spatial analysis of groundwater recharge across the catchments revealed key differences between the model configurations. Configuration BL struggled to simulate recharge rates exceeding 250 mm yr^{-1}, despite such values being common in the study area. However, it performed well in catchments with low recharge values, consistently producing lower recharge estimates compared to GW and GW-RC.

For configurations GW and GW-RC, groundwater recharge rates were influenced by catchment size and total precipitation. Larger catchments with higher precipitation exhibited greater recharge, while smaller, drier catchments showed lower recharge rates. This relationship indicates that these configurations better capture broad spatial trends in groundwater recharge compared to configuration BL, which showed less sensitivity to variations in precipitation and catchment size. Furthermore, GW and GW-RC displayed similar spatial patterns. Configuration GW exhibited the highest variability between catchments, whereas GW-RC produced estimates of average annual recharge that were more consistent with PACES data across most catchments. Future studies should further investigate how spatial characteristics of catchments affect the overall dynamics of hydrological variables in this context.

Supporting these observations, Chemingui *et* al. (2015) found the average recharge rates across different seasons at three locations in the “des Anglais” catchment. The numbers retrieve in their work closely align with those simulated by the GW-RC configuration: winter (58 vs 50 mm), spring (58 vs 54 mm), summer (92 vs 60 mm), and fall (52 vs 72 mm).

Furthermore, Rivard *et* al. (2014) utilized the HELP infiltration model to simulate recharge for a catchment in Eastern Canada, reporting average recharge rates of 67 mm in winter, 62 mm in spring, 27 mm in summer, and 76 mm in fall. These findings align with our results from configuration GW-RC, which also show peak recharge occurring in fall rather than in spring, differentiating it from the other configurations. Configuration GW aligns less precisely with these specific seasonal

patterns, with a peak recharge in spring, but still outperforms BL in terms of matching the documented recharge rates from PACES.

Recharge rates from GW-RC compare favorably with observed seasonal fluctuations in the literature. Overall, GW-RC's alignment with empirical data and its ability to simulate hydrological processes more accurately make it a preferable model configuration for studying and predicting hydrological dynamics under varied climatic conditions.

In this study, the GW-RC configuration demonstrated that assigning a minor weight to recharge in the objective function can significantly enhance WaSiM's capability to represent hydrological variables accurately, even with non-exact prior recharge data. This approach underscores, again, the potential of leveraging prior information to refine model outputs, suggesting that even a modest emphasis on recharge within the calibration framework can lead to substantial improvements in model realism. This finding is particularly noteworthy as it implies that effective model calibration does not necessarily require precise initial recharge estimates if the calibration process is appropriately managed. It also points to the broader applicability of using informed yet flexible calibration strategies to improve hydrological models under varied conditions, highlighting a path forward for enhancing model accuracy with limited prior data.

4.4 Practical implications, general applicability and limitations

This research has practical applications beyond hydrological modeling. Integrating groundwater recharge into model calibration, as demonstrated in the GW-RC configuration, offers a more comprehensive approach to representing key hydrological variables. This approach is particularly valuable for improving predictions of water resources under varying climate conditions, as it enhances the accuracy of inputs critical to models of forest growth (Ford *et* al., 2011; Grant *et* al., 2013). As climate change continues to alter hydrological dynamics, the reliance on physically based models becomes crucial. These models are favored over conceptual ones or even machine learning based models because they can be adapted more readily to varying conditions, ensuring more robust predictions under climate change scenarios. For example, a strong recent trend is the use of deep learning architectures in hydrological modelling (Kratzert et al. 2018, 2019; Arsenault et al. 2023). These models simulate streamflow with generally better accuracy than traditional hydrological models, but they lack any mechanism to investigate internal and intermediate hydrological variables. Such adaptability is also critical for effective water resource management and mitigation of climate impacts (Wilby, 2005; Ludwig *et* al., 2009; Poulin *et* al., 2011). By improving the representation of hydrological processes, the GW-RC configuration may enhance the model's ability to simulate hydrological responses under changing climatic conditions. This is especially important given the non-stationarity of climate, where historical hydrological relationships no longer hold under future conditions. In this context, calibrating models using physically meaningful constraints, such as groundwater recharge, may improve their ability to capture shifting hydrological patterns and enhance confidence in assessments of climate change impacts on hydrological variables.

This research emphasizes the need to calibrate hydrological models using not only streamflow but also other variables such as groundwater recharge. This approach aligns with findings from other studies such as Yassin *et* al. (2017) and Dembélé *et* al. (2020), which advocate for multi-objective calibrations that enhance model reliability across different hydrological

variables. By integrating measurements from diverse sources such as satellite data and in-situ measurements, models can avoid the pitfalls of calibration based solely on streamflow, which might not capture the full spectrum of watershed dynamics. Bouaziz *et* al. (2021) further illustrate this point by showing that hydrological models calibrated solely on streamflow can produce differing results when validated against other hydrological variables. This highlights the risk of equifinality, where different parameter sets yield similar streamflow outputs but diverge for other hydrological processes. Without proper constraints, such as incorporating groundwater recharge into calibration, models may generate realistic streamflow simulations while misrepresenting key internal processes. This issue is evident in configurations BL and GW, which fail to accurately capture certain underlying hydrological dynamics.

The methodology developed in this study has broad applicability beyond the specific context of Southern Québec. This approach can be valuable in a variety of geographic regions and hydrological settings, given similar contexts of equifinality (i.e. more processes and parameters than the data can support). Moreover, this multi-variable calibration method can enhance the accuracy of other distributed hydrological models by improving the representation of groundwater recharge related processes. Similar calibration techniques using remote-sensing data have been applied successfully in different settings, demonstrating that incorporating additional hydrological variables in calibration improves model performance.

Nevertheless, it is crucial to address the limitations of this study. The models' performance in replicating hydrological processes like soil frost impacts and its implications on runoff and recharge remain unknown. Future studies would benefit from incorporating field measurements alongside a broader range of climatic and hydrological conditions. Expanding the research to include different geographic regions with similar soil and climate characteristics could significantly enhance the validation and applicability of the findings.

Additionally, the selected catchments in this study range from 525 km² to 6,840 km², which may limit the generalizability of the findings to catchments outside this size range. Future research could investigate smaller or larger catchments to determine whether the observed trends and calibration impacts remain consistent across different watershed scales.

Furthermore, the choice of objective function presents another limitation. This study primarily relied on the Kling-Gupta Efficiency (KGE) for streamflow calibration. However, alternative metrics such as SPAtial EFficiency (SPAEF) (Koch *et* al., 2018) could enable a more comprehensive evaluation of multiple hydrological components when using distributed hydrological models. The lack of sufficient spatially distributed observations prevented the application of SPAEF in this study, but future research could explore its use, particularly in conjunction with remote sensing data to better assess the spatial coherence of hydrological variables.

Moreover, the uncertainty inherent in modeling, especially with configurations that involve complex interactions of multiple variables, poses a continuous challenge. The study's reliance on specific data sets like PACES also introduces potential biases that could influence the generalizability of the findings. The two-step calibration adopted for GW-RC, necessitated by the absence of high-quality, spatially distributed recharge observations, limits the extent to which a fully direct comparison with GW can be achieved. Future work with access to such datasets could implement a single-step calibration using both streamflow and recharge, enabling a more controlled assessment of the effects of internal recharge constraints. It's essential

for future research to explore these limitations, perhaps by expanding the range of observational data used for model validation.

In terms of practical implementations and further research, continuing to refine the calibration of hydrological models to include diverse hydrological variables can enhance their utility in real-world applications. Such efforts will help in developing more accurate flood forecasting models, improving water resource management strategies, and crafting more effective climate adaptation measures for forest, agricultural and anthropogenic ecosystems. This study advances calibration techniques in hydrological modeling, but further work is needed to develop universally reliable models.

5 Conclusion

This study examined the nuances of hydrological modeling under different calibration settings using WaSiM model across 34 catchments classified under climate zones Dfb and Dfc in Eastern North America. By implementing three distinct model configurations, BL (baseline model), GW (physical groundwater model), and GW-RC (physical groundwater and recharge calibration model), this research has demonstrated that incorporating groundwater recharge alongside streamflow during calibration process leads to a representation of hydrological processes that better aligns with expected system behavior.

The results indicate that the GW-RC configuration, enhanced with groundwater recharge calibration, aligns more closely with estimated groundwater recharge rates, thereby providing a more precise representation of groundwater behaviour both spatially and seasonally. The study also underscores the importance of extending calibration beyond traditional streamflow metrics to include other hydrological variables like groundwater recharge. This approach helps to mitigate the risks of equifinality.

Given the successful application of these methodologies within Eastern North American catchments, it presents an intriguing premise for their applicability to other geographical areas with similar hydrological contexts. Further research could explore how these calibration techniques perform under different hydrological conditions, potentially broadening our understanding of these relationships.

Appendix A

Table A1. Kling-Gupta efficiency values across studied catchments during calibration and validation periods, for the three calibrations configurations. Each row corresponds to a specific catchment, identified by its basin number.

Catchment		Calibration			Validation		
Code	**Name**	**BL**	**GW**	**GW-RC**	**BL**	**GW**	**GW-RC**
1	Bonaventure	0.835	0.849	0.817	0.897	0.892	0.861
2	York	0.847	0.872	0.815	0.843	0.889	0.814
3	Dartmouth	0.888	0.882	0.842	0.899	0.907	0.828
4	Matane	0.906	0.901	0.877	0.908	0.900	0.861
5	Rimouski	0.920	0.905	0.870	0.831	0.812	0.827
6	Des Trois-Pistoles	0.898	0.895	0.848	0.783	0.754	0.712
7	Du Loup	0.872	0.852	0.800	0.795	0.749	0.696
8	Ouelle	0.900	0.896	0.834	0.838	0.846	0.792
9	Famine	0.826	0.814	0.754	0.794	0.798	0.745
10	Bécancour	0.861	0.859	0.788	0.820	0.816	0.765
11	Nicolet Sud-Ouest	0.828	0.810	0.771	0.801	0.770	0.746
12	Nicolet	0.804	0.799	0.744	0.811	0.792	0.767
13	Eaton	0.769	0.768	0.637	0.738	0.741	0.661
14	Au Saumon	0.836	0.815	0.717	0.790	0.774	0.713
15	Noire	0.823	0.813	0.723	0.767	0.770	0.694
16	Rouge	0.830	0.842	0.798	0.838	0.829	0.844
17	Gatineau	0.817	0.840	0.796	0.807	0.831	0.772
18	Kinojévis	0.765	0.850	0.784	0.695	0.711	0.735
19	Mattawin	0.852	0.814	0.740	0.799	0.758	0.751
20	Croche	0.835	0.835	0.833	0.839	0.840	0.831
21	Vermillon	0.835	0.853	0.747	0.808	0.809	0.733
22	Batiscan	0.878	0.856	0.801	0.884	0.847	0.796
23	Sainte-Anne	0.872	0.860	0.833	0.852	0.847	0.829
24	Bras du Nord	0.853	0.864	0.856	0.859	0.863	0.869
25	Ouareau	0.855	0.881	0.818	0.839	0.837	0.765
26	L'Assomption	0.865	0.886	0.851	0.829	0.859	0.821
27	De l'Achigan	0.869	0.851	0.829	0.700	0.720	0.701
28	Du Loup	0.808	0.783	0.800	0.786	0.721	0.753
29	Petit Saguenay	0.895	0.879	0.843	0.864	0.857	0.800
30	Petite rivière Péribonca	0.833	0.876	0.775	0.819	0.839	0.742
31	Métabetchouane	0.872	0.861	0.806	0.801	0.769	0.670
32	Valin	0.880	0.882	0.826	0.842	0.888	0.793
33	Sainte-Marguerite Nord-Est	0.872	0.854	0.810	0.854	0.833	0.772
34	Godbout	0.857	0.864	0.799	0.838	0.861	0.777

Table A2. Multiple streamflow metrics values during calibration (2000-2009) and validation (1990-1999) periods, for the three configurations.

Metric		Calibration			Validation		
		BL	**GW**	**GW-RC**	**BL**	**GW**	**GW-RC**
KGE	μ	0.852	0.852	0.799	0.816	0.820	0.772
	σ	0.034	0.036	0.050	0.055	0.049	0.056
Pearson Coefficient	μ	0.855	0.855	0.804	0.844	0.845	0.797
	σ	0.034	0.036	0.050	0.040	0.039	0.049
Bias ratio	μ	0.998	0.990	0.985	1.030	1.024	1.018
	σ	0.021	0.024	0.023	0.054	0.052	0.050
Variability ratio	μ	0.996	1.005	1.020	1.022	1.028	1.055
	σ	0.023	0.013	0.027	0.083	0.075	0.079
NSE	μ	0.704	0.706	0.603	0.677	0.679	0.558
	σ	0.059	0.076	0.091	0.091	0.073	0.120
RMSE	μ	20.926	20.749	24.317	22.877	22.621	26.545
	σ	11.018	10.681	13.055	12.594	11.665	13.870
Percent bias	μ	0.155	0.074	1.263	-3.563	-2.760	-1.756
	σ	1.399	2.209	2.132	5.792	5.628	5.650
MAE	μ	11.364	11.198	13.287	12.444	12.045	14.432
	σ	6.621	6.259	8.106	7.966	7.043	8.638

690

Appendix B

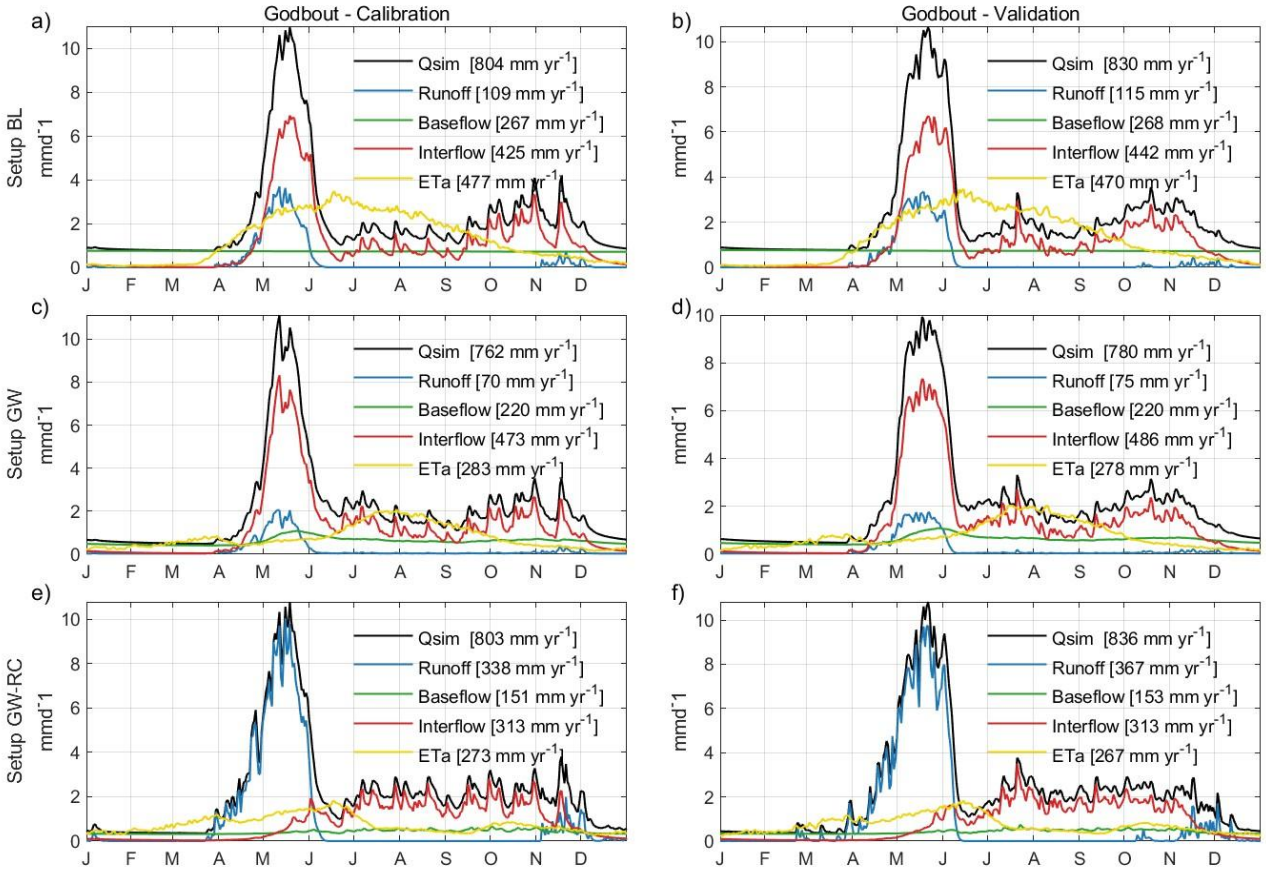
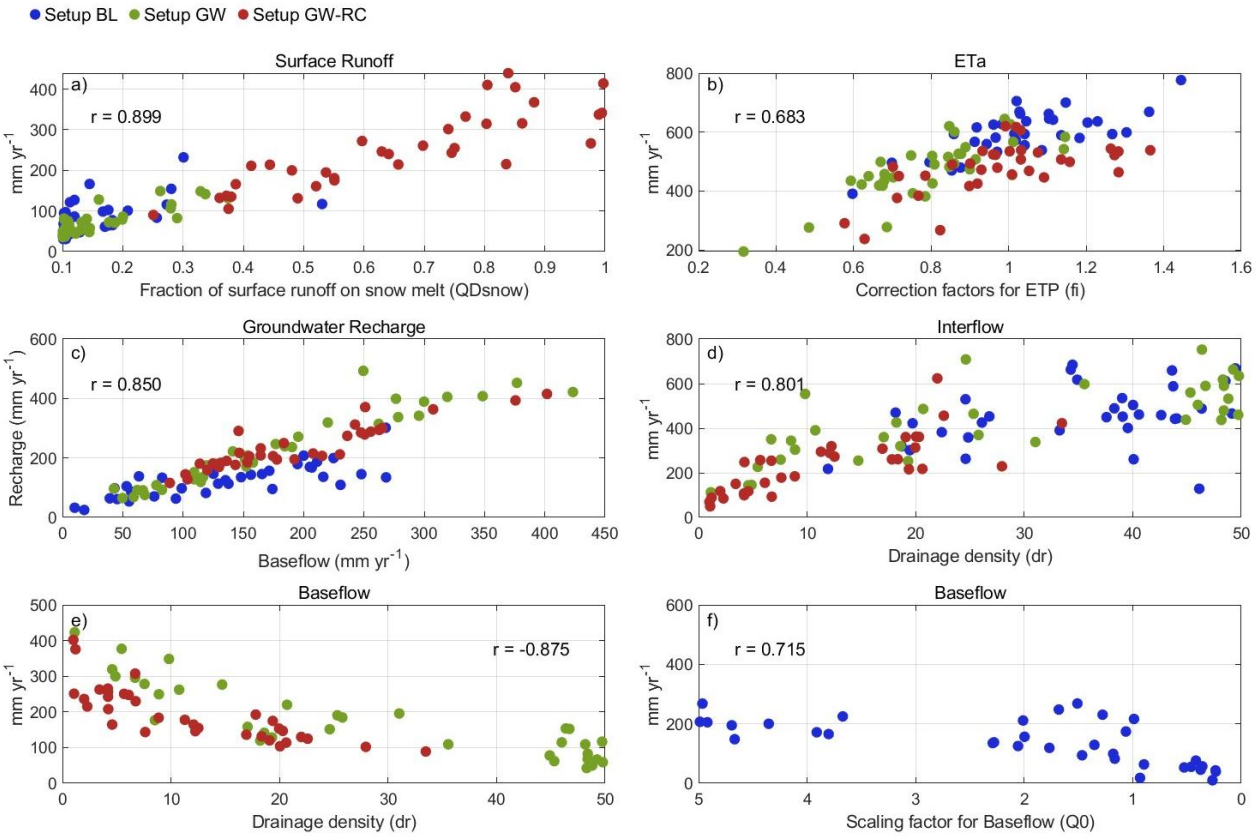


Figure B1. Detailed hydrological variable hydrograph for Godbout catchment during both the calibration and validation phases and for the three configurations. Calibration results are shown in panels (a), (c), and (e) for Configurations BL, GW, and GW-RC, respectively, while validation results are depicted in panels (b), (d), and (f). These hydrographs demonstrate how baseflow, interflow and runoff contribute to total streamflow throughout the year, with noted annual totals provided for a comprehensive comparison.

Appendix C



700 Figure C1. Correlations between key hydrological variables and calibration parameters for three model configurations.

Appendix D

A.

Catchment		k_D			k_H			d_r			QD_{Snow}			c_0			K_{rec}			T_0			$T_{R/S}$			C_{WH}		
Code	**Name**	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C
1	Bonaventure	24.7	24.7	25.0	25.0	24.8	24.5	38	46	19	0.1	0.1	0.4	1.8	2.2	2.6	0.5	0.1	0.9	-0.6	-0.1	0.4	0.3	0.5	0.0	0.1	0.1	0.1
2	York	25.0	24.8	24.7	24.1	24.3	14.9	38	18	5	0.1	0.1	0.8	1.2	1.8	1.4	0.6	0.4	0.3	-1.3	-0.2	0.3	1.3	0.2	1.6	0.1	0.1	0.1
3	Dartmouth	21.6	24.4	24.7	20.9	17.7	1.8	49	50	8	0.1	0.1	0.8	2.1	2.1	2.1	0.8	0.9	0.2	-0.3	-0.1	0.1	1.4	1.0	2.3	0.1	0.1	0.2
4	Matane	24.7	24.9	24.6	25.0	24.7	19.8	44	48	20	0.1	0.1	0.6	2.3	2.9	2.5	0.1	0.2	0.2	-0.1	0.6	0.6	1.8	0.7	1.1	0.1	0.1	0.1
5	Rimouski	24.8	24.9	24.6	24.9	25.0	23.8	43	17	4	0.2	0.3	0.4	1.5	1.9	2.1	0.2	0.2	0.9	-0.1	0.6	-0.3	0.3	-2.4	2.5	0.2	0.2	0.2
6	Des Trois-Pistoles	24.7	24.7	25.0	24.9	24.5	1.6	49	48	2	0.2	0.2	0.8	1.9	2.8	2.1	1.0	1.0	0.5	0.2	1.0	0.7	-3.1	-2.6	-2.3	0.3	0.2	0.2
7	Du Loup	24.7	23.4	24.9	24.3	24.1	16.8	41	8	4	0.1	0.2	0.6	2.0	2.6	1.9	0.2	0.2	0.9	-0.8	0.7	0.6	1.7	0.3	-2.6	0.2	0.3	0.3
8	Ouelle	24.7	24.9	24.6	24.6	23.6	3.5	40	50	19	0.1	0.1	0.6	2.3	2.5	2.3	0.1	0.1	1.0	-0.1	0.0	0.6	1.7	1.8	1.2	0.3	0.3	0.3
9	Famine	22.6	24.5	25.0	19.8	11.8	1.6	46	48	18	0.1	0.1	0.7	1.9	2.7	2.6	1.0	0.7	0.3	-1.1	-0.8	0.7	1.1	1.2	0.3	0.2	0.3	0.2
10	Bécancour	24.9	24.4	24.7	23.6	25.0	23.5	35	47	20	0.1	0.1	0.7	2.7	2.5	2.0	0.1	0.9	0.9	0.2	0.2	0.0	1.3	1.3	1.7	0.1	0.1	0.2
11	Nicolet Sud-Ouest	25.0	24.1	24.9	24.9	24.7	16.4	49	36	33	0.1	0.1	0.7	2.9	2.8	2.5	0.6	1.0	0.9	0.7	0.5	0.8	-3.0	-0.2	-2.9	0.1	0.2	0.2
12	Nicolet	24.9	24.7	24.8	22.8	19.1	5.0	25	45	28	0.1	0.2	1.0	3.0	2.0	1.7	0.9	0.9	1.0	0.7	-0.6	0.3	-2.9	0.2	0.0	0.1	0.2	0.2
13	Eaton	24.7	24.3	20.9	3.5	2.9	2.0	20	49	9	0.3	0.4	1.0	2.8	3.0	1.8	0.9	0.9	0.2	0.0	-0.1	-0.2	-3.6	-0.3	-3.8	0.1	0.2	0.1
14	Au Saumon	25.0	24.7	24.9	20.6	12.7	2.7	49	49	18	0.1	0.3	1.0	2.5	3.0	2.1	0.9	1.0	0.2	-0.8	0.0	0.1	1.7	-2.6	-1.5	0.1	0.1	0.1
15	Noire	25.0	24.5	24.9	13.8	14.9	2.5	27	45	21	0.1	0.3	1.0	2.1	3.0	2.2	0.8	0.8	1.0	-1.0	0.3	0.3	0.5	-1.2	0.4	0.3	0.3	0.2
16	Rouge	17.8	23.5	24.3	24.8	24.7	24.8	40	5	6	0.1	0.1	0.3	0.8	1.7	0.9	0.1	0.9	0.2	-2.3	-0.8	-3.2	0.1	0.7	1.9	0.1	0.2	0.3
17	Gatineau	24.1	22.6	24.9	12.7	24.3	24.8	19	5	6	0.3	0.3	0.6	0.9	1.7	0.9	0.3	0.2	0.7	-3.0	-0.1	0.0	3.1	-3.8	-1.9	0.3	0.3	0.2
18	Kinojévis	24.9	22.9	24.0	24.4	24.2	21.4	46	19	7	0.5	0.1	0.5	1.1	1.8	0.9	0.8	0.3	0.8	0.9	0.5	-1.9	-3.2	-0.9	3.5	0.3	0.1	0.2
19	Mattawin	22.6	24.1	24.7	24.8	24.9	20.9	25	15	4	0.1	0.1	0.6	1.7	2.2	0.8	0.1	0.1	0.6	-0.3	1.0	-1.3	0.8	-1.7	1.3	0.2	0.1	0.2
20	Croche	9.3	11.1	24.8	21.6	13.7	24.8	25	1	17	0.1	0.1	0.4	0.8	2.2	1.0	0.3	0.3	0.2	-1.2	1.3	-0.2	-0.5	-2.2	-3.2	0.3	0.2	0.2
21	Vermillon	12.7	12.2	24.7	20.1	21.5	23.9	12	5	1	0.2	0.1	0.5	1.0	1.9	1.2	0.3	0.3	0.3	-0.5	0.5	0.5	-3.2	0.9	0.7	0.3	0.2	0.1
22	Batiscan	22.6	22.5	24.9	24.5	24.8	1.8	18	25	1	0.2	0.1	0.8	0.7	0.5	0.7	0.2	0.2	0.6	-1.9	-2.1	0.2	-1.8	-3.0	-1.7	0.2	0.3	0.2
23	Sainte-Anne	24.9	24.2	24.9	23.6	18.5	7.3	34	25	23	0.3	0.3	0.8	1.7	1.4	1.0	0.2	0.3	0.9	0.3	0.0	-0.1	-3.2	-0.2	-0.4	0.1	0.1	0.1
24	Bras du Nord	24.7	24.9	25.0	24.4	24.3	24.2	34	46	22	0.1	0.1	0.4	1.4	1.0	1.3	0.7	0.2	0.7	0.4	-1.5	0.3	-3.4	-1.2	-0.7	0.1	0.2	0.1
25	Ouareau	21.8	23.0	24.7	24.8	16.3	24.7	39	7	12	0.2	0.1	0.5	1.6	2.5	1.6	0.1	0.1	0.6	0.5	0.8	0.7	-1.2	0.5	-1.9	0.2	0.2	0.1
26	L'Assomption	24.7	22.5	24.8	24.3	24.5	25.0	33	31	13	0.1	0.1	0.4	1.1	2.3	1.6	0.1	0.1	0.7	-0.8	0.3	0.7	0.8	-1.0	-0.2	0.1	0.1	0.2
27	De l'Achigan	24.9	23.6	24.6	24.7	1.6	6.0	44	48	11	0.1	0.1	0.4	2.2	2.8	2.7	0.2	0.6	0.8	0.3	0.6	0.6	-2.8	-0.2	-2.1	0.1	0.2	0.2
28	Du Loup	24.9	23.4	24.5	24.9	24.7	24.3	19	19	3	0.2	0.1	0.4	1.0	1.3	1.6	0.4	0.2	0.1	-0.8	0.6	0.3	0.0	-0.4	2.8	0.2	0.2	0.2
29	Petit Saguenay	24.9	24.6	24.8	24.8	3.6	2.7	22	9	2	0.1	0.2	0.7	1.5	1.8	1.0	0.1	0.3	0.9	0.2	0.0	-0.1	0.3	-3.8	1.3	0.1	0.2	0.2
30	Petite rivière Péribonca	24.7	24.3	24.8	23.6	24.4	1.2	40	26	4	0.1	0.1	0.8	1.7	0.7	0.5	0.3	0.3	0.7	1.5	-3.5	-0.8	-3.1	2.4	0.2	0.1	0.3	0.3
31	Métabetchouane	24.2	24.8	25.0	24.4	18.0	16.5	26	9	1	0.1	0.2	0.5	0.6	1.4	1.1	0.1	0.1	0.5	-2.2	-1.1	1.3	-0.4	-0.8	-2.5	0.3	0.3	0.2
32	Valin	25.0	22.8	25.0	23.5	19.4	17.2	44	10	7	0.2	0.1	0.9	1.2	2.4	0.8	0.1	0.5	0.9	0.7	1.4	0.4	-2.5	1.2	0.3	0.1	0.1	0.2
33	Sainte-Marguerite Nord-Est	25.0	24.7	24.6	24.3	17.5	18.2	39	11	12	0.1	0.3	0.9	1.7	2.4	1.3	0.1	0.3	0.9	0.6	1.1	0.7	-2.6	-3.5	-0.9	0.1	0.1	0.1
34	Godbout	24.8	23.2	24.8	23.8	24.7	24.6	44	21	20	0.3	0.1	0.9	1.3	2.4	1.4	1.0	0.3	0.6	0.1	0.6	0.5	0.7	1.0	-1.8	0.2	0.2	0.1
	Average	**23.4**	**23.3**	**24.6**	**22.6**	**19.9**	**14.6**	**35**	**28**	**12**	**0.2**	**0.2**	**0.7**	**1.7**	**2.1**	**1.6**	**0.4**	**0.4**	**0.6**	**-0.4**	**0.0**	**0.1**	**-0.6**	**-0.5**	**-0.2**	**0.2**	**0.2**	**0.2**

B.

Catchment		C_{rfr}			$f_{i,summer}$			$f_{i,fall}$			$f_{i,winter}$			$f_{i,spring}$			Kol			K_{XY}			d_z		
Code	**Name**	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C
1	Bonaventure	0.7	0.2	0.7	1.3	1.1	1.0	0.5	0.2	0.6	0.5	2.0	1.5	1.4	0.8	0.7	-	83	9	-	3.2	1.4	1.1	1.3	1.4
2	York	0.6	0.6	0.8	1.7	1.1	0.5	1.0	0.6	0.2	0.2	0.2	1.3	1.6	1.4	1.8	-	5	15	-	0.6	1.5	1.1	1.3	1.4
3	Dartmouth	0.6	0.9	0.8	1.2	1.2	0.9	1.4	0.7	0.7	0.3	0.2	0.9	0.6	0.7	0.6	-	35	55	-	0.2	1.7	0.9	1.0	1.3
4	Matane	0.6	0.6	0.9	1.2	1.6	1.3	1.7	0.8	0.9	0.4	1.0	1.0	0.9	0.2	0.5	-	97	44	-	3.2	1.6	1.1	1.1	0.9
5	Rimouski	0.6	0.6	0.7	1.2	1.2	0.9	1.7	0.8	0.7	0.3	0.8	0.2	0.9	0.7	1.1	-	23	35	-	0.3	0.5	1.1	1.1	0.8
6	Des Trois-Pistoles	0.7	0.4	0.4	1.3	1.3	1.2	1.2	1.5	1.1	0.1	0.3	0.6	1.0	0.6	0.7	-	44	56	-	1.5	2.4	1.4	0.8	0.8
7	Du Loup	0.7	1.0	0.7	1.8	1.2	1.1	1.4	0.6	0.8	0.4	0.1	0.8	0.8	1.1	1.0	-	20	41	-	0.2	2.3	1.1	1.3	1.2
8	Ouelle	0.7	0.8	1.0	1.1	1.1	1.1	2.0	1.5	1.0	0.7	1.0	1.2	1.1	0.9	0.9	-	40	92	-	0.2	2.2	1.1	1.0	1.2
9	Famine	0.6	0.7	0.7	1.0	1.4	1.0	1.7	0.1	0.8	0.2	0.5	0.9	1.1	0.7	0.7	-	21	63	-	0.2	0.6	1.2	0.9	1.1
10	Bécancour	0.6	1.0	0.6	1.5	1.3	1.2	0.6	1.5	0.7	0.7	0.7	1.4	0.9	0.7	0.5	-	47	75	-	1.4	3.6	1.1	1.4	1.3
11	Nicolet Sud-Ouest	0.7	0.3	0.9	1.2	1.4	1.2	1.0	0.5	0.9	0.8	0.6	1.1	1.1	0.9	0.9	-	62	97	-	0.2	0.6	1.1	1.0	0.9
12	Nicolet	0.8	1.0	0.8	1.2	1.7	1.7	1.0	0.6	0.4	0.9	0.2	1.3	1.0	0.9	0.6	-	65	62	-	0.2	3.9	1.3	0.8	1.1
13	Eaton	0.6	0.8	0.9	0.9	1.2	0.8	1.2	1.4	1.3	0.8	0.4	1.1	1.6	1.0	0.9	-	21	37	-	0.2	2.8	1.2	0.8	1.2
14	Au Saumon	0.8	0.9	0.9	1.0	1.1	0.9	1.3	0.7	0.7	0.9	0.1	0.9	0.7	0.6	0.6	-	20	66	-	0.2	3.9	1.3	0.8	1.2
15	Noire	0.8	0.7	0.7	1.2	1.2	1.0	1.1	1.5	1.6	0.6	0.2	0.7	1.2	1.0	0.9	-	64	94	-	0.3	2.6	1.2	0.8	1.1
16	Rouge	1.0	0.9	1.0	1.2	1.1	1.7	0.9	0.9	0.6	1.5	0.3	1.7	1.0	0.5	0.2	-	16	25	-	2.2	2.6	1.2	1.3	0.9
17	Gatineau	0.6	0.8	0.8	1.3	1.2	0.8	0.4	0.5	1.2	1.9	0.1	1.3	1.1	0.6	0.9	-	15	14	-	2.2	4.0	1.0	1.3	1.1
18	Kinojévis	0.7	0.5	0.5	0.4	1.0	0.7	1.2	1.0	1.8	1.0	0.9	1.3	1.7	0.2	0.6	-	4	21	-	3.2	3.9	1.4	1.1	1.2
19	Mattawin	0.6	0.7	0.4	1.2	0.6	0.9	1.4	1.3	1.7	0.4	0.3	1.7	1.1	1.0	0.8	-	14	21	-	3.0	3.8	1.4	1.4	1.2
20	Croche	0.6	0.7	1.0	1.7	1.2	1.0	1.3	0.2	1.6	1.4	0.1	1.3	0.6	0.8	0.7	-	26	29	-	0.8	3.6	0.9	0.8	1.1
21	Vermillon	0.7	0.5	0.7	1.9	1.4	1.0	0.5	1.0	2.0	0.5	0.3	2.0	1.0	0.9	0.2	-	4	21	-	1.0	4.0	1.3	1.3	1.1
22	Batiscan	0.7	0.7	0.7	1.6	1.2	0.4	1.4	0.7	1.4	0.2	0.5	1.9	1.0	1.2	1.3	-	21	32	-	4.0	3.5	1.1	1.3	1.0
23	Sainte-Anne	0.7	0.8	0.9	1.1	1.0	0.9	0.6	0.3	0.6	0.4	0.6	1.6	1.0	0.8	0.6	-	7	40	-	1.3	2.9	1.2	1.1	1.2
24	Bras du Nord	0.8	0.8	0.8	1.5	1.2	0.9	1.0	0.5	0.5	0.7	0.2	1.9	1.0	0.9	0.7	-	3	33	-	1.7	3.1	0.8	1.3	0.9
25	Ouareau	0.6	0.9	0.8	1.4	1.0	1.0	1.8	0.6	0.7	1.5	0.5	1.0	0.5	0.7	0.1	-	26	29	-	1.0	3.9	0.9	1.1	1.2
26	L'Assomption	0.8	0.8	0.3	1.6	1.3	1.3	0.6	0.7	1.3	0.4	0.7	1.3	1.3	0.6	0.5	-	9	12	-	1.2	3.2	1.3	1.4	1.2
27	De l'Achigan	0.7	0.7	0.7	1.2	1.3	1.4	0.6	0.4	0.4	0.7	0.9	0.6	0.9	0.5	0.4	-	34	32	-	0.4	1.2	1.1	1.1	0.8
28	Du Loup	0.9	0.6	1.0	1.6	1.2	1.2	1.0	0.8	0.4	0.4	0.5	1.5	1.4	0.9	0.8	-	4	16	-	4.0	2.6	1.1	1.2	1.0
29	Petit Saguenay	0.7	0.1	0.3	1.7	1.4	0.6	1.9	0.9	1.2	0.5	0.3	1.3	1.3	0.8	1.5	-	25	46	-	1.8	1.8	1.1	1.3	1.2
30	Petite rivière Péribonca	1.0	1.0	0.4	1.7	1.1	0.4	1.9	0.8	1.6	0.3	1.7	2.0	1.9	0.9	1.5	-	6	35	-	3.7	4.0	1.3	1.1	1.3
31	Métabetchouane	0.7	0.9	0.7	1.5	1.2	0.7	1.0	1.2	1.7	1.2	0.4	1.6	1.2	0.2	1.2	-	3	28	-	3.2	3.8	1.1	1.2	1.0
32	Valin	0.6	0.7	0.8	1.4	0.7	0.1	0.2	0.7	0.3	0.3	0.2	0.7	0.9	0.4	1.2	-	12	77	-	2.2	3.9	1.1	1.1	1.2
33	Sainte-Marguerite Nord-Est	0.7	0.8	0.4	0.9	0.6	0.1	0.4	0.3	0.6	0.1	0.2	1.0	0.9	0.3	0.8	-	27	80	-	1.3	2.8	1.1	1.3	1.2
34	Godbout	0.7	0.6	0.9	1.0	0.7	0.2	0.8	0.5	0.9	0.3	1.2	1.5	1.4	0.3	0.7	-	5	34	-	2.5	3.3	1.4	1.2	1.2
	Average	**0.7**	**0.7**	**0.7**	**1.3**	**1.2**	**0.9**	**1.1**	**0.8**	**1.0**	**0.6**	**0.5**	**1.2**	**1.1**	**0.7**	**0.8**	**-**	**27**	**43**	**-**	**1.5**	**2.7**	**1.1**	**1.1**	**1.1**

Figure D1. Final calibrated parameter values for all catchments across each model configuration (BL, GW, GW-RC).

Code and data availability

The calibrated WaSiM model for all configurations discussed in this study is publicly accessible at https://osf.io/h9rsj/ (Talbot et al., 2024). This dataset encompasses control files, input parameters and output files from both calibration and validation phases.

Author contribution

FT, JDS and RA conceptualized the study's methodology and designed the model configurations. FT carried out the model simulations, data collection and analysis, interpretation of results and led the writing of the manuscript. JDS and RA supervised the project. AP, GD, JDS, RA and FT discussed the results and contributed to the final manuscript.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

This work was funded jointly by the ministère des Ressources naturelles et des Forêts (Quebec, Canada, project number 112332187 conducted at the Direction de la recherche forestière and led by Jean-Daniel Sylvain) and the Forest research service contract number 3322-2022-2187-01 obtained by Richard Arsenault from the Ministère des Ressources naturelles et des Forêts (Quebec, Canada). The authors also acknowledge the use of ChatGPT-4 for assistance in correcting spelling mistakes and improving the flow of text during the manuscript preparation process. The base map in Fig. 1 was created using ArcGIS® software by Esri. ArcGIS® and ArcMap™ are the intellectual property of Esri and are used herein under license. Copyright © Esri. All rights reserved. For more information about Esri® software, please visit www.esri.com.

References

Acero Triana, J. S., Chu, M. L., Guzman, J. A., Moriasi, D. N., and Steiner, J. L.: Beyond model metrics: The perils of calibrating hydrologic models, Journal of Hydrology, 578, 124032, https://doi.org/10.1016/j.jhydrol.2019.124032, 2019.

Akumu, C. E., Woods, M., Johnson, J. A., Pitt, D. G., Uhlig, P., and McMurray, S.: GIS-fuzzy logic technique in modeling soil depth classes: Using parts of the Clay Belt and Hornepayne region in Ontario, Canada as a case study, Geoderma, 283, 78–87, https://doi.org/10.1016/j.geoderma.2016.07.028, 2016.

Arsenault, R., Poulin, A., Côté, P., and Brissette, F.: Comparison of Stochastic Optimization Algorithms in Hydrological Model Calibration, Journal of Hydrologic Engineering, 19, 1374–1384, https://doi.org/10.1061/(ASCE)HE.1943-5584.0000938, 2014.

Arsenault, R., Martel, J.-L., Brunet, F., Brissette, F., and Mai, J.: Continuous streamflow prediction in ungauged basins: long short-term memory neural networks clearly outperform traditional hydrological models, Hydrol. Earth Syst. Sci., 27, 139–157, https://doi.org/10.5194/hess-27-139-2023, 2023.

Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., and Wood, E. F.: Present and future Köppen-Geiger climate classification maps at 1-km resolution, Sci Data, 5, 180214, https://doi.org/10.1038/sdata.2018.214, 2018.

Bormann, H. and Elfert, S.: Application of WaSiM-ETH model to Northern German lowland catchments: model performance in relation to catchment characteristics and sensitivity to land use change, in: Advances in Geosciences, 1–10, https://doi.org/10.5194/adgeo-27-1-2010, 2010.

Bouaziz, L. J. E., Fenicia, F., Thirel, G., de Boer-Euser, T., Buitink, J., Brauer, C. C., De Niel, J., Dewals, B. J., Drogue, G., Grelier, B., Melsen, L. A., Moustakas, S., Nossent, J., Pereira, F., Sprokkereef, E., Stam, J., Weerts, A. H., Willems, P., Savenije, H. H. G., and Hrachowitz, M.: Behind the scenes of streamflow model performance, Hydrology and Earth System Sciences, 25, 1069–1095, https://doi.org/10.5194/hess-25-1069-2021, 2021.

Boumaiza, L., Walter, J., Chesnaux, R., Lambert, M., Jha, M. K., Wanke, H., Brookfield, A., Batelaan, O., Galvão, P., Laftouhi, N.-E., and Stumpp, C.: Groundwater recharge over the past 100 years: Regional spatiotemporal assessment and climate change impact over the Saguenay-Lac-Saint-Jean region, Canada, Hydrological Processes, 36, e14526, https://doi.org/10.1002/hyp.14526, 2022.

Buffin-Bélanger, T., Chaillou, G., Cloutier, C.-A., Touchette, M., Hétu, B., and McCormack, R.: Programme d'acquisition de connaissance sur les eaux souterraines du nord-est du Bas-Saint-Laurent (PACES-NEBSL): Rapport final, Université du Québec à Rimouski, Département de biologie, chimie et géographie, Rimouski, Québec, 2015.

Carrier, M.-A., Lefebvre, R., Rivard, C., Parent, M., Ballard, J.-M., Benoit, N., Vigneault, H., Beaudry, C., Malet, X., Laurencelle, M., Gosselin, J.-S., Ladevèze, P., Thériault, R., Beaudin, I., Michaud, A., Pugin, A., Morin, R., Crow, H., Gloaguen, E., Bleser, J., Martin, A., and Lavoie, D.: Portrait des ressources en eau souterraine en Montérégie Est, Québec, Canada, Institut national de la recherche scientifique (INRS), Commission géologique du Canada (CGC), Organisme de bassin versant de la Yamaska (OBV Yamaska), Institut de recherche et développement en agroenvironnement (IRDA), Québec, 2013.

Chemingui, A., Sulis, M., and Paniconi, C.: An assessment of recharge estimates from stream and well data and from a coupled surface-water/groundwater model for the des Anglais catchment, Quebec (Canada), Hydrogeol J, 23, 1731–1743, https://doi.org/10.1007/s10040-015-1299-1, 2015.

Chen, J., Sudicky, E., Gula, J., Peltier, W., Park, Y., and Ross, M.: Impact of Climate Change on Canadian Surface Water and Groundwater Resources: A Continental-scale Hydrologic Modelling Study using Multiple High-resolution RCM Projections, 2011.

Cloutier, V., Blanchette, D., Dallaire, P.-L., Nadeau, S., Rosa, E., and Roy, M.: Projet d'acquisition de connaissances sur les eaux souterraines de l'Abitibi-Témiscamingue (partie 1), Groupe de recherche sur l'eau souterraine, Institut de recherche en mines et en environnement, Université du Québec en Abitibi-Témiscamingue, Québec, 2013.

Cloutier, V., Rosa, E., Nadeau, S., Dallaire, P.-L., Blanchette, D., and Roy, M.: Projet d'acquisition de connaissances sur les eaux souterraines de l'Abitibi-Témiscamingue (partie 2), Groupe de recherche sur l'eau souterraine, Institut de recherche en mines et en environnement, Université du Québec en Abitibi-Témiscamingue, Québec, 2015.

Comeau, G., Talbot-Poulin, M. C., Tremblay, Y., Ayotte, S., Molson, J., Lemieux, J. M., Montcoudiol, N., Therrien, R., Fortier, R., Therrien, P., and Fabien-Ouellet, G.: Projet d'acquisition de connaissances sur les eaux souterraines en Outaouais, Département de géologie et de génie géologique, Université Laval, Québec, 2013.

2015 Land Cover of North America at 30 meters: http://www.cec.org/north-american-environmental-atlas/land-cover-30m-2015-landsat-and-rapideye/, last access: 1 December 2023.

Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., Wehberg, J., Wichmann, V., and Böhner, J.: System for Automated Geoscientific Analyses (SAGA) v. 2.1.4, Geoscientific Model Development, 8, 1991–2007, https://doi.org/10.5194/gmd-8-1991-2015, 2015.

Croteau, A., Nastev, M., and Lefebvre, R.: Groundwater Recharge Assessment in the Chateauguay River Watershed, Canadian Water Resources Journal / Revue canadienne des ressources hydriques, 35, 451–468, https://doi.org/10.4296/cwrj3504451, 2010.

Cunningham, S. C., Thomson, J. R., Mac Nally, R., Read, J., and Baker, P. J.: Groundwater change forecasts widespread forest dieback across an extensive floodplain system: Groundwater change forecasts forest dieback, Freshwater Biology, 56, 1494–1508, https://doi.org/10.1111/j.1365-2427.2011.02585.x, 2011.

Dembélé, M., Hrachowitz, M., Savenije, H. H. G., Mariéthoz, G., and Schaefli, B.: Improving the Predictive Skill of a Distributed Hydrological Model by Calibration on Spatial Patterns With Multiple Satellite Data Sets, Water Resources Research, 56, e2019WR026085, https://doi.org/10.1029/2019WR026085, 2020.

Dubois, E., Larocque, M., Gagné, S., and Meyzonnat, G.: Simulation of long-term spatiotemporal variations in regional-scale groundwater recharge: Contributions of a water budget approach in southern Quebec, https://doi.org/10.5194/hess-2021-71, 26 February 2021.

Duethmann, D., Anderson, M., Maneta, M. P., and Tetzlaff, D.: Improving process-consistency of an ecohydrological model through inclusion of spatial patterns of satellite-derived land surface temperature, Journal of Hydrology, 628, 130433, https://doi.org/10.1016/j.jhydrol.2023.130433, 2024.

Farjad, B., Gupta, A., and Marceau, D. J.: Annual and Seasonal Variations of Hydrological Processes Under Climate Change Scenarios in Two Sub-Catchments of a Complex Watershed, Water Resour Manage, 30, 2851–2865, https://doi.org/10.1007/s11269-016-1329-3, 2016.

Ficchì, A., Perrin, C., and Andréassian, V.: Hydrological modelling at multiple sub-daily time steps: Model improvement via flux-matching, Journal of Hydrology, 575, 1308–1327, https://doi.org/10.1016/j.jhydrol.2019.05.084, 2019.

Ford, C. R., Laseter, S. H., Swank, W. T., and Vose, J. M.: Can forest management be used to sustain water-based ecosystem services in the face of climate change?, Ecological Applications, 21, 2049–2067, https://doi.org/10.1890/10-2246.1, 2011.

Förster, K., Meißl, G., Marke, T., Pohl, S., Garvelmann, J., Schulla, J., and Strasser, U.: A new snow-vegetation interaction extension for the Water Balance Simulation Model (WaSiM), 2620, 2017.

Förster, K., Garvelmann, J., Meißl, G., and Strasser, U.: Modelling forest snow processes with a new version of WaSiM, Hydrological Sciences Journal, 63, 1540–1557, https://doi.org/10.1080/02626667.2018.1518626, 2018.

Gädeke, A., Hölzel, H., Koch, H., Pohle, I., and Grünewald, U.: Analysis of uncertainties in the hydrological response of a model-based climate change impact assessment in a subcatchment of the Spree River, Germany, Hydrological Processes, 28, 3978–3998, https://doi.org/10.1002/hyp.9933, 2014.

Van Genuchten, M. Th.: A Closed-form Equation for Predicting the Hydraulic Conductivity of Unsaturated Soils, Soil Science Soc of Amer J, 44, 892–898, https://doi.org/10.2136/sssaj1980.03615995004400050002x, 1980.

Géobase du réseau hydrographique du Québec (GRHQ) - Données Québec: https://www.donneesquebec.ca/recherche/dataset/grhq, last access: 30 November 2023.

Grant, G. E., Tague, C. L., and Allen, C. D.: Watering the forest for the trees: an emerging priority for managing water in forest landscapes, Frontiers in Ecology and the Environment, 11, 314–321, https://doi.org/10.1890/120209, 2013.

Hamon, W. R.: Estimating Potential Evapotranspiration, Transactions of the American Society of Civil Engineers, 128, 324–338, https://doi.org/10.1061/TACEAT.0008673, 1963.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Hock, R.: Temperature index melt modelling in mountain areas, Journal of Hydrology, 282, 104–115, https://doi.org/10.1016/S0022-1694(03)00257-9, 2003.

Imhoff, R. O., van Verseveld, W. J., van Osnabrugge, B., and Weerts, A. H.: Scaling Point-Scale (Pedo)transfer Functions to Seamless Large-Domain Parameter Estimates for High-Resolution Distributed Hydrologic Modeling: An Example for the Rhine River, Water Resources Research, 56, e2019WR026807, https://doi.org/10.1029/2019WR026807, 2020.

Jacobs, J. M.: Ecohydrology: Darwinian Expression of Vegetation Form and Function, EoS Transactions, 84, 345–345, https://doi.org/10.1029/2003EO350009, 2003.

Jasper, K., Calanca, P., and Fuhrer, J.: Changes in summertime soil water patterns in complex terrain due to climatic change, Journal of Hydrology, 327, 550–563, https://doi.org/10.1016/j.jhydrol.2005.11.061, 2006.

Jeong, S., Hong, M., and Song, J.: Soil depth estimation in mountainous areas by using GIS and satellite images, Landslides, 19, 2711–2726, https://doi.org/10.1007/s10346-022-01919-2, 2022.

Kirchner, J. W.: Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology, Water Resources Research, 42, 2005WR004362, https://doi.org/10.1029/2005WR004362, 2006.

Kling, H., Fuchs, M., and Paulin, M.: Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios, Journal of Hydrology, 424–425, 264–277, https://doi.org/10.1016/j.jhydrol.2012.01.011, 2012.

Koch, J., Demirel, M. C., and Stisen, S.: The SPAtial EFficiency metric (SPAEF): multiple-component evaluation of spatial patterns for optimization of hydrological models, Geoscientific Model Development, 11, 1873–1886, https://doi.org/10.5194/gmd-11-1873-2018, 2018.

Kratzert, F., Klotz, D., Brenner, C., Schulz, K., and Herrnegger, M.: Rainfall–runoff modelling using Long Short-Term Memory (LSTM) networks, Hydrol. Earth Syst. Sci., 22, 6005–6022, https://doi.org/10.5194/hess-22-6005-2018, 2018.

Kratzert, F., Klotz, D., Herrnegger, M., Sampson, A. K., Hochreiter, S., and Nearing, G. S.: Toward Improved Predictions in Ungauged Basins: Exploiting the Power of Machine Learning, Water Resources Research, 55, 11344–11354, https://doi.org/10.1029/2019WR026065, 2019.

Larocque, M., Gagné, S., Tremblay, L., and Meyzonnat, G.: Projet de connaissance des eaux souterraines du bassin versant de la rivière Bécancour et de la MRC de Bécancour - Rapport final, Ministère du Développement durable, de l'Environnement, de la Faune et des Parcs, Québec, 2013.

Larocque, M., Gagné, S., Barnetche, D., Meyzonnat, G., Graveline, M.-H., and Ouellet, M. A.: Projet de connaissance des eaux souterraines du bassin versant de la zone Nicolet et de la partie basse de la zone Saint-François - Rapport final, Ministère du Développement durable, de l'Environnement et de la Lutte contre les changements climatiques, Québec, 2015.

Larocque, M., Levison, J., Martin, A., and Chaumont, D.: A review of simulated climate change impacts on groundwater resources in Eastern Canada, Canadian Water Resources Journal / Revue canadienne des ressources hydriques, 44, 22–41, https://doi.org/10.1080/07011784.2018.1503066, 2019.

Latifovic, R., Homer, C., Ressl, R., Pouliot, D. A., Hossian, S., Colditz, R., Olthof, I., Chandra, G., and Victoria, A.: North American Land Change Monitoring System, Remote Sensing of Land Use and Land Cover: Principles and Applications, 303–324, https://doi.org/10.1201/b11964-24, 2012.

Lefebvre, R., Ballard, J.-M., Vigneault, H., Beaudry, C., Berthot, L., Légaré-Couture, G., Parent, M., Laurencelle, M., Malet, X., Therrien, A., Michaud, A., Desjardins, J., Drouin, A., Cloutier, M. H., Grenier, J., Bourgault, M.-A., Larocque, M., Pellerin, S., Graveline, M.-H., Janos, D., and Molson, J.: Portrait des ressources en eau souterraine en Chaudière-Appalaches, Québec, Canada, Institut national de la recherche scientifique (INRS), Institut de recherche et développement en agroenvironnement (IRDA), Regroupement des organismes de bassins versants de la Chaudière-Appalaches (OBV-CA), Québec, 2015.

Lighthill, M. J. and Whitham, G. B.: On kinematic waves I. Flood movement in long rivers, Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences, 229, 281–316, https://doi.org/10.1098/rspa.1955.0088, 1955.

De Lima Ferreira, P. M. and Da Paz, A. R.: Enhanced calibration of a distributed hydrological model in the Brazilian Semi-Arid: integrating spatiotemporal evapotranspiration and streamflow data, Environ Earth Sci, 83, 345, https://doi.org/10.1007/s12665-024-11663-6, 2024.

Liu, X., Yang, K., Ferreira, V. G., and Bai, P.: Hydrologic Model Calibration With Remote Sensing Data Products in Global Large Basins, Water Resources Research, 58, e2022WR032929, https://doi.org/10.1029/2022WR032929, 2022.

Ludwig, R., May, I., Turcotte, R., Vescovi, L., Braun, M., Cyr, J.-F., Fortin, L.-G., Chaumont, D., Biner, S., Chartier, I., Caya, D., and Mauser, W.: The role of hydrological model complexity and uncertainty in climate change impact assessment, Advances in Geosciences, 21, 63–71, https://doi.org/10.5194/adgeo-21-63-2009, 2009.

MacMillan, R. A., Pettapiece, W. W., Nolan, S. C., and Goddard, T. W.: A generic procedure for automatically segmenting landforms into landform elements using DEMs, heuristic rules and fuzzy logic, Fuzzy Sets and Systems, 113, 81–109, https://doi.org/10.1016/S0165-0114(99)00014-7, 2000.

Maitre, D. L., Scott, D., and Colvin, C.: A review of information on interactions between vegetation and groundwater, Water SA, 1999.

Markhali, S. P., Poulin, A., and Boucher, M.-A.: Spatio-temporal discretization uncertainty of distributed hydrological models, Hydrological Processes, 36, e14635, https://doi.org/10.1002/hyp.14635, 2022.

Mei, Y., Mai, J., Do, H. X., Gronewold, A., Reeves, H., Eberts, S., Niswonger, R., Regan, R. S., and Hunt, R. J.: Can Hydrological Models Benefit From Using Global Soil Moisture, Evapotranspiration, and Runoff Products as Calibration Targets?, Water Resources Research, 59, e2022WR032064, https://doi.org/10.1029/2022WR032064, 2023.

Meyer Oliveira, A., Fleischmann, A. S., and Paiva, R. C. D.: On the contribution of remote sensing-based calibration to model hydrological and hydraulic processes in tropical regions, Journal of Hydrology, 597, 126184, https://doi.org/10.1016/j.jhydrol.2021.126184, 2021.

Ministère du Développement durable, de l'Environnement et de la Lutte contre les changements climatiques: Hydroclimatic Atlas of Southern Québec (3rd ed.), available at: https://www.cehq.gouv.qc.ca/atlas-hydroclimatique/index-en.htm, last access: 30 November 2023.

Mizukami, N., Clark, M. P., Newman, A. J., Wood, A. W., Gutmann, E. D., Nijssen, B., Rakovec, O., and Samaniego, L.: Towards seamless large-domain parameter estimation for hydrologic models, Water Resources Research, 53, 8020–8040, https://doi.org/10.1002/2017WR020401, 2017.

Natkhin, M., Steidl, J., Dietrich, O., Dannowski, R., and Lischeid, G.: Differentiating between climate effects and forest growth dynamics effects on decreasing groundwater recharge in a lowland region in Northeast Germany, Journal of Hydrology, 448–449, 245–254, https://doi.org/10.1016/j.jhydrol.2012.05.005, 2012.

Orellana, F., Verma, P., Loheide, S. P., and Daly, E.: Monitoring and modeling water-vegetation interactions in groundwater-dependent ecosystems, Reviews of Geophysics, 50, 2011RG000383, https://doi.org/10.1029/2011RG000383, 2012.

Persaud, E., Levison, J., MacRitchie, S., Berg, S. J., Erler, A. R., Parker, B., and Sudicky, E.: Integrated modelling to assess climate change impacts on groundwater and surface water in the Great Lakes Basin using diverse climate forcing, Journal of Hydrology, 584, 124682, https://doi.org/10.1016/j.jhydrol.2020.124682, 2020.

Peters-Lidard, C. D., Clark, M., Samaniego, L., Verhoest, N. E. C., Van Emmerik, T., Uijlenhoet, R., Achieng, K., Franz, T. E., and Woods, R.: Scaling, similarity, and the fourth paradigm for hydrology, Hydrol. Earth Syst. Sci., 21, 3701–3713, https://doi.org/10.5194/hess-21-3701-2017, 2017.

Pool, S., Fowler, K., and Peel, M.: Benefit of Multivariate Model Calibration for Different Climatic Regions, Water Resources Research, 60, e2023WR036364, https://doi.org/10.1029/2023WR036364, 2024.

Poulin, A., Brissette, F., Leconte, R., Arsenault, R., and Malo, J.-S.: Uncertainty of hydrological modelling in climate change impact studies in a Canadian, snow-dominated river basin, Journal of Hydrology, 409, 626–636, https://doi.org/10.1016/j.jhydrol.2011.08.057, 2011.

Pradhan, A. and Indu, J.: Uncertainty in Calibration of Variable Infiltration Capacity Model, 89–108, https://doi.org/10.1007/978-3-030-02197-9_4, 2019.

Rajib, A., Evenson, G. R., Golden, H. E., and Lane, C. R.: Hydrologic model predictability improves with spatially explicit calibration using remotely sensed evapotranspiration and biophysical parameters, Journal of Hydrology, 567, 668–683, https://doi.org/10.1016/j.jhydrol.2018.10.024, 2018.

Richards, L. A.: Capillary conduction of liquids through porous mediums, Physics, 1, 318–333, https://doi.org/10.1063/1.1745010, 1931.

Rivard, C., Paniconi, C., Vigneault, H., and Chaumont, D.: A watershed-scale study of climate change impacts on groundwater recharge (Annapolis Valley, Nova Scotia, Canada), Hydrological Sciences Journal, 59, 1437–1456, https://doi.org/10.1080/02626667.2014.887203, 2014.

Rößler, O. and Löffler, J.: Potentials and limitations of modelling spatio-temporal patterns of soil moisture in a high mountain catchment using WaSiM-ETH, Hydrological Processes, 24, 2182–2196, https://doi.org/10.1002/hyp.7663, 2010.

Rössler, O., Diekkrüger, B., and Löffler, J.: Potential drought stress in a Swiss mountain catchment—Ensemble forecasting of high mountain soil moisture reveals a drastic decrease, despite major uncertainties, Water Resources Research, 48, 2011WR011188, https://doi.org/10.1029/2011WR011188, 2012.

Rouleau, A., Daigneault, R., Walter, J., Germaneau, D., Tremblay, M.-L., Lambert, M., Dallaire, C., Chesnaux, R., Cousineau, P. A., Roy, D. W., Noël, D., Simard, P., Moisan, A., and Thouvignon, N.: Résultats du programme d'acquisition de connaissances sur les eaux souterraines de la région Saguenay-Lac-Saint-Jean, Centre d'études sur les ressources minérales, Université du Québec à Chicoutimi, Chicoutimi, Québec, 2013.

Samaniego, L., Kumar, R., and Attinger, S.: Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale, Water Resources Research, 46, 2008WR007327, https://doi.org/10.1029/2008WR007327, 2010.

Samaniego, L., Kumar, R., Thober, S., Rakovec, O., Zink, M., Wanders, N., Eisner, S., Müller Schmied, H., Sutanudjaja, E. H., Warrach-Sagi, K., and Attinger, S.: Toward seamless hydrologic predictions across spatial scales, Hydrol. Earth Syst. Sci., 21, 4323–4346, https://doi.org/10.5194/hess-21-4323-2017, 2017.

Schäfer, C., Fäth, J., Kneisel, C., Baumhauer, R., and Ullmann, T.: Multidimensional hydrological modeling of a forested catchment in a German low mountain range using a modular runoff and water balance model, Front. For. Glob. Change, 6, https://doi.org/10.3389/ffgc.2023.1186304, 2023.

SIIGSOL-100m - Carte des propriétés du sol: https://www.donneesquebec.ca/recherche/dataset/siigsol-100m-carte-des-proprietes-du-sol, last access: 1 December 2023.

Shepard, D.: A two-dimensional interpolation function for irregularly-spaced data, in: Proceedings of the 1968 23rd ACM national conference, New York, NY, USA, 517–524, https://doi.org/10.1145/800186.810616, 1968.

Sun, G., Wei, X., Hao, L., Sanchis, M. G., Hou, Y., Yousefpour, R., Tang, R., and Zhang, Z.: Forest hydrology modeling tools for watershed management: A review, Forest Ecology and Management, 530, 120755, https://doi.org/10.1016/j.foreco.2022.120755, 2023.

Sylvain, J.-D., Anctil, F., and Thiffault, E.: Using bias correction and ensemble modelling for predictive mapping and related uncertainty: A case study in digital soil mapping, Geoderma, 403, 1–29, https://doi.org/10.1016/j.geoderma.2021.115153, 2021.

T. W. Chu and A. Shirmohammadi: Evaluation of the SWAT models hydrology component in the Piedmont physiographic region of Maryland, Transactions of the ASAE, 47, 1057–1073, https://doi.org/10.13031/2013.16579, 2004.

Talbot, F., Sylvain, J.-D., Drolet, G., Poulin, A., and Arsenault, R.: HESS paper data - Wasim Calibration with internal state variable constraints, https://doi.org/10.17605/OSF.IO/H9RSJ, 2024.

Tarek, M., Brissette, F. P., and Arsenault, R.: Evaluation of the ERA5 reanalysis as a potential reference dataset for hydrological modelling over North America, Hydrology and Earth System Sciences, 24, 2527–2544, https://doi.org/10.5194/hess-24-2527-2020, 2020.

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

Valencia Giraldo, M. del C., Ricard, S., and Anctil, F.: Assessment of the Potential Hydrological Impacts of Climate Change in Quebec—Canada, a Refined Neutral Approach, Water, 15, 584, https://doi.org/10.3390/w15030584, 2023.

Vose, J. M., Sun, G., Ford, C. R., Bredemeier, M., Otsuki, K., Wei, X., Zhang, Z., and Zhang, L.: Forest ecohydrological research in the 21st century: what are the critical needs?, Ecohydrology, 4, 146–158, https://doi.org/10.1002/eco.193, 2011.

Wang, Q., Deng, H., and Jian, J.: Hydrological Processes under Climate Change and Human Activities: Status and Challenges, Water, 15, 4164, https://doi.org/10.3390/w15234164, 2023.

Weil, R. and Brady, N.: The Nature and Properties of Soils. 15th edition, 2017.

Wilby, R. L.: Uncertainty in water resource model parameters used for climate change impact assessment, Hydrological Processes, 19, 3201–3219, https://doi.org/10.1002/hyp.5819, 2005.

Xu, C., Widén, E., and Halldin, S.: Modelling hydrological consequences of climate change—Progress and challenges, Adv. Atmos. Sci., 22, 789–797, https://doi.org/10.1007/BF02918679, 2005.

Yassin, F., Razavi, S., Wheater, H., Sapriza-Azuri, G., Davison, B., and Pietroniro, A.: Enhanced identification of a hydrologic model using streamflow and satellite water storage data: A multicriteria sensitivity analysis and optimization approach, Hydrological Processes, 31, 3320–3333, https://doi.org/10.1002/hyp.11267, 2017.

WaSiM-ETH Documentation: http://www.wasim.ch/en/products/wasim_description.htm, last access: 30 November 2023.