Examining the challenges of simulating surface water – groundwater interactions in a post-glacial environment

Félix Turgeon[1], Marie Larocque[1*], Guillaume Meyzonnat[1], Sarah Dorner[2], Marc-André Bourgault[1]

[1] Département des sciences de la Terre et de l'atmosphère & GEOTOP Research Center, Université du Québec à Montréal, Montréal, Québec H3C 3P8, 514-987-3000, larocque.marie@uqam.ca

[2] Département des génies civil, géologique et des mines, École Polytechnique de Montréal, Québec H3T 1J7

For the use of the editors

Paper #:

Submitted on:

Accepted on:

Application - Research – Commentary – Book Review:

Abstract

Although integrated models are increasingly used for water management purposes, detailed applications of these models under different conditions are necessary to guide their implementation. The objective of this study was to examine some of the challenges encountered when simulating surface water – groundwater interactions in a post-glacial geological environment. The integrated MikeSHE model was used to simulate transient-state heads and flows in the Raquette River watershed in the Vaudreuil-Soulanges region of southwestern Quebec (Canada) over a two year period. This application benefited from a detailed hydrogeological database recently developed for the region. Overall, flows, heads, and groundwater inputs to the river were adequately simulated. A sensitivity analysis has shown that many hydrogeologic and surface flow parameters have an impact on both flow rates and heads, thus underlining the importance of using an integrated model to study watershed-scale water issues. Additional flow rate measurements to improve the quality of rating curves and continuous flow measurements in tributaries, could improve model calibration. An explicit simulation of unsaturated zone infiltration processes, including soil flow, plant, and evaporation processes, as well as the inclusion of the agricultural tile drainage system, could reduce simulation errors. Extending the model calibration over a longer period, including contrasting hydrological conditions, would make the model more robust in view of its use for water management under land use and climate change conditions. Nevertheless, this work demonstrated that, using data readily available for southern Québec aquifers, it is possible to build an integrated model that is representative of actual hydrological conditions. The maintenance and improvement of this model for long-term use is recommended.

Résumé

Les modèles entièrement couplés sont de plus en plus utilisés pour réaliser la gestion intégrée des ressources en eau, mais les exemples détaillés d'applications sont encore peu nombreux. L'objectif de cette recherche était d'examiner les défis posés par la simulation des interactions eaux de surface – eaux souterraines dans un environnement géologique post-glaciaire. Le modèle couplé MikeSHE a été utilisé pour simuler les charges et les débits sur le bassin de la rivière à la Raquette dans la région de Vaudreuil-Soulanges (Québec, Canada), sur une période de deux années. Ce travail a bénéficié d'une base de données hydrogéologiques récemment développée pour la région. Les résultats montrent que les débits, les charges et les flux échangés entre l'aquifère et la rivière sont relativement bien simulés, ce qui indique que le modèle est utile dans son état actuel, malgré le fait que certaines incertitudes et limitations aient été identifiés. L'analyse de sensibilité a montré que plusieurs paramètres hydrogéologiques et d'écoulement superficiel ont un impact à la fois sur les débits et les charges, ce qui met en évidence l'importance d'utiliser un modèle couplé pour la gestion intégrée de l'eau. Des mesures de débits aux tributaires pour toute l'année ainsi que des courbes de tarages basées sur une plus large gamme de débits pourraient améliorer le calage du modèle. Une représentation explicite de la zone non saturée, incluant les processus d'écoulement dans le sol, le prélèvement par les plantes et l'évaporation, de même que l'inclusion du réseau de drains agricoles, pourraient contribuer à réduire les erreurs de simulation. Le prolongement de la période de calibration pour inclure des conditions hydrologiques contrastées contribuerait à rendre le modèle plus robuste pour des applications liées à la

gestion de l'eau en conditions de changement dans l'utilisation du territoire et de changements climatiques. Néanmoins, cette recherche a montré que les données utilisées, qui correspondent à celles généralement disponibles dans le sud du Québec, permettent de construire un modèle couplé qui représente bien les conditions hydrologiques actuelles. Il est recommandé de poursuivre le développement du modèle pour des applications à long terme.

Keywords: Groundwater, surface water, MikeSHE model, Raquette River, southern Québec (Canada)

Introduction

Over the last two decades, surface water and groundwater reservoirs have increasingly been studied as two expressions of a single resource, a paradigm described by Winter et al. (1998). In humid climates, aquifers sustain water inflow to rivers, maintaining riparian ecosystems (Brunke and Gonser 1997; Hayashi and Rosenberry 2002) and thermal refuges for fish species (Kurylyk et al. 2015), especially during low flow periods. Increasing groundwater pumping (Konikow and Kendy 2005), and the resulting water table drawdown, can reduce groundwater inflow into rivers (McCallum et al. 2013) and induce river water infiltration into the riverbed (Chen and Shu, 2002). Poor groundwater quality can also affect river water quality (Rozemeijer and Broers 2007).

As a result of the growing recognition of their coupled nature, surface water – groundwater interactions have been monitored and simulated in a large range of geological and climatic conditions throughout the world. Reviews of related processes and methods are reported in Fleckenstein et al. (2010) and Kalbus et al. (2006). Integrated model applications are increasingly used to understand the complex interactions between different water reservoirs, and are increasingly used as tools to guide integrated water management by consultants and regulatory agencies. However, applications of integrated models are not yet routine, due to either the lack of available data, or the lack of human and computer resources to develop the models. A thorough analysis of integrated flow conditions, including a sensitivity analysis and an examination of the areas needing improvement to build a more robust model, is usually reserved to academic contexts. As a result, integrated models that are used and improved upon over the long-term to guide integrated water management are still rare. Such models are necessary to address low

flow, as well as flooding conditions, while allowing the different water users to perform their various activities.

The challenge of simulating surface and groundwater flows together is heightened when studying complex 3D geological environments, where topography, stratigraphy, and hydraulic parameters of the media can vary substantially over short distances. Topography and geology have large effects on surface water – groundwater interactions (e.g., Gleeson and Manning 2008). Complex post-glacial environments are especially difficult to assess, due to the high horizontal spatial and vertical heterogeneity of the geological units that typically often comprise them. In addition, the unconsolidated sediments often lie on fractured bedrock, where the spatial heterogeneity of hydraulic properties can be very large. These difficulties can be exacerbated in small agricultural rivers that are prone to large variations in flow rates as a result of the flashy behavior due to the limited extent of permeable areas that allow infiltration relative to the large clayey areas, where tile drainage is pervasive. Such conditions are particularly difficult to simulate using integrated flow models, because they require a good characterization of both runoff and groundwater flow processes to represent the full range of rapidly changing flow conditions.

The objective of this study was to examine some of the challenges encountered when simulating surface water – groundwater interactions in a post-glacial environment. It is intended to contribute to the development of models that will become truly useful water management tools. This study makes use of recently updated geological, hydrogeological, and hydrological data for the Vaudreuil-Soulanges region of southern Quebec, Canada (Larocque et al. 2015), obtained through a detailed regional aquifer characterization

project. This watershed is an example of post-glacial geological conditions in which surface water – groundwater interactions are not easily assessed, and as such is an excellent case study area for which to build an integrated model.

Study area

Physiography and land use

The Raquette River watershed (133 km^2) is located in the St. Lawrence Lowlands, 50 km west of Montreal, in the province of Quebec (Canada) (Figure 1). Approximately 75% of the area is covered by Light Detection and Ranging (LiDAR) data (GéoMont 2011) with a vertical accuracy of 15 cm on a horizontal grid of 1 m^2. Over the rest of the study area, topographic data are derived from a digital elevation model with a vertical precision of 3 m and a spatial resolution of 10 m. The watershed has a marked topography over short distances, with elevations ranging from 23 m (all elevations stated as meters above sea level) near the Outaouais River to 228 m at the peak of Mount Rigaud. Between the Saint-Lazare Hill and Mount Rigaud, a northeast-southwest aligned bedrock depression corresponds to the Sainte-Marthe corridor. At the base of Mount Rigaud, the agricultural plain ranges from 75 m to 60 m in elevation at the southern boundary of the Raquette watershed. At the northern boundary, topography ranges from 23 to 35 m. The sandy Hudson Hill is located along the northeast boundary, reaches 80 m, and corresponds to the Raquette River water divide (Larocque et al. 2015).

Figure 1. Location of the Raquette River watershed in the Vaudreuil-Soulanges region of southern Quebec (Canada), including topography and monitoring stations.

Land use is classified as agricultural for more than half of the watershed (53%) (Larocque et al. 2015). Agricultural fields are mainly located upgradient in the watershed, south of Mount Rigaud. Forest covers 41% of the watershed, mostly on Mount Rigaud, and on the Saint-Lazare and Sainte-Justine-de-Newton hills. Urban or residential areas cover 5% of the watershed, and the remaining 2% is occupied by power lines. There are very few wetlands and no lakes capable of water retention in the watershed.

Geology

Bedrock geology in the study area is composed of a sedimentary sequence of Paleozoic age with low deformation. It overlays the Precambrian bedrock and corresponds to the Grenville Province (Hofmann 1972). This sedimentary bedrock is located at the base of the sedimentary sequence of the St. Lawrence Lowlands.

The bottom of the sedimentary sequence is composed of sandstone of the Upper Cambrian Potsdam Group, with a total thickness of up to 450 m (Williams et al. 2010). The two formations of the Potsdam Group, Covey Hill and Cairnside, are composed of fossil-free feldspar sandstone and fossiliferous quartzite sandstone. The Lower Ordovician Beekmantown Group overlies the Potsdam sandstone, and is composed of dolomites, limestones, and calcareous sandstones, approximately 250 m thick, including the Theresa and Beauharnois formations (Hofmann 1972). The Chazy Group rests on the Beekmantown Group, and is marked by a contribution of detrital elements during the Middle Ordovician. The Chazy Group is 100 m thick, and includes the members of the Ste-Thérèse Formation and of the Laval Formation.

The higher topography of Mount Rigaud results from an intrusion in the Grenville bedrock, which occurred at the end of the Cambrian and the beginning of the Ordovician (Greig 1968). The sedimentary sandstone bedrock outcrops in certain locations along the Raquette River bed over a few hundred meters.

Surface deposits mapped in the study area were emplaced during the last glacial age, the late Wisconsinan (Roy and Godbout 2014; Figure 2a). They show contrasting conditions, with the top of Mount Rigaud covered by thin and discontinuous till, and the flanks of Mount Rigaud having till deposits of up to 15 m thick. Some till deposits are also visible to the southwest of Mount Rigaud. The Champlain Sea clay deposits are found everywhere in the plain and can reach 45 m in thickness. Major sand deposits are also present in the Raquette River watershed on the Hudson and Saint-Lazare hills, as well as in the area of Sainte-Justine-de-Newton. Between Mount Rigaud and the Hudson Hill, the Raquette River meanders on silty sand deposits. Littoral sands of glacio-marine origin are found in the Saint-Lazare area and on the slopes of Mount Rigaud. Available drilling data (Technorem 2005; MDDELCC 2013) and seismic information (Hobson and Tremblay 1962) provide evidence that buried channels of sand and gravel could locally reach 40 m thick below 20 to 30 m of Champlain Sea clay deposits. No storage coefficients or specific storage coefficients were available for the study area.

Figure 2. Geological and hydrogeological conditions in the Raquette River watershed; a) Quaternary deposits (modified from Roy and Godbout 2014), and b) piezometric map (modified from Larocque et al. 2015).

Hydrology

The Raquette River is a fourth order river, which drains surface and groundwater over a distance of 34 km. Upstream, it begins on the till crests located southwest of Mount Rigaud. It drains eastward through the agricultural plain, contours Mount Rigaud on its eastern flank, and discharges into the Outaouais River. The watershed has three additional tributaries, which join the Raquette River upstream of the Sainte-Marthe valley, as well as other small streams also flowing from Mount Rigaud into the river. In the agricultural area of the watershed, a substantial drainage network is in place (ditches and drains).

Water levels were monitored at three locations on the Raquette River between March 2013 and fall 2014 by Larocque et al. (2015; see Figure 1 for the locations of gauging stations). Station 1 was installed in mid-March 2013, and water levels were measured throughout the year (water level measured at 30 minute intervals with an ultrasonic *Hobo Onset* logger). However, winter values were discarded because of ice cover between January and early-April 2014). Stations 2 and 3 were installed in mid-July and mid-May 2013 respectively (water level measured at 30 minute intervals with *Solinst Levelogger*s in perforated plastic tubes at the river bed). These probes were removed between December 2013 and mid-May 2014 to prevent damage from ice. Water levels were transformed into flow rates using rating curves and velocities measured at different times during the study period (with an HACH portable Doppler flow meter (GENEQ) for low flows, and a *StreamPro Acoustic Doppler Current Profiler* (DASCO) for high flows). Two rating curves were constructed for station 3, due to equipment failure in December 2013. Twenty-three flow rate measurements were taken at station 1 between May 2013 and August 2014 ($RMSE_{stat1} = 2.3\ m^3s^{-1}$), 18 measurements were taken at

station 2 between August 2013 and August 2014 ($RMSE_{stat2} = 0.2\ m^3\ s^{-1}$), and 26 measurements were taken at station 3 between April 2013 and August 2014 ($RMSE_{stat3} = 0.1\ m^3\ s^{-1}$ in 2013 and $0.2\ m^3\ s^{-1}$ in 2014). For all three stations, and particularly for station 1, the uncertainty is greatest for high flow rates. Detailed river cross-sections at the three gauging stations were obtained using a differential GPS. The average daily flow rates at gauging station 1, located downgradient in the study area, closest to the Raquette River outlet, varied between 0.4 and 88 $m^3\ s^{-1}$. It is clear from the rating curve RMSE at this station that a large error is associated with the high flow value. Flow rates at gauging station 2 varied between 0.1 and 17.4 $m^3.s^{-1}$, while those at gauging station 3 varied between 0.03 and 24.3 $m^3\ s^{-1}$.

The evolution of flow rates along the Raquette River flow path was measured twice in August 2014 (August 12 and 25; no precipitation for five consecutive days prior to measurements) at 23 locations between stations 1 and 3 using the velocity and cross-section method. The flow rates at the outlet of the three main tributaries of the Raquette River were measured in the same way on each occasion. The measured flow rates for the two dates were very similar, with differences within the uncertainty usually associated with this type of flow measurement (from 2 to 20%; Carter and Anderson 1963; Pelletier 1988; Sauer and Meyer 1992; Harmel et al. 2006), and varied between 0.008 and 0.220 $m^3\ s^{-1}$ (3). Upgradient, where the river flows on clay deposits in an agricultural setting (between points 1 and 8), the flow rate is low and increases slowly. The Raquette River occasionally runs dry in this area. In the Sainte-Marthe Valley (between points 8 and 9), there is a marked increase in flow rates, even though the river flows on clay deposits. The increase in flow rates is smaller, but relatively constant, for

these two dates between stations 9 and 19. There is a drop in flow rates on both dates between stations 19 and 20, a further increase between stations 20 and 22, and a decrease in flow rates in the lowest reaches of the river.

Figure 3. Measured flow rates at 23 locations along the Raquette River on August12th and 25th 2015. The black arrows indicate tributaries 1 (T1), 2 (T2), and 3 (T3). The grey arrows indicate the main streams. The dashed vertical lines indicate the locations of gauging stations 1, 2, and 3.

The three main tributaries reach the Raquette River between measurement points 6 and 8. In August 2014, the flow rates were very low for all three tributaries. At their outlets, the flow rates were 1) constant, at 0.008 $m^3 s^{-1}$, for tributary 1 (T1), 2) varied between 0.002 and 0.003 $m^3 s^{-1}$ for tributary 2 (T2), and 3) varied between 0.002 and 0.004 $m^3 s^{-1}$ for tributary 3 (T3).

Hydrogeology

Available hydraulic conductivities were reported in Larocque et al. (2015), and are derived from three observation wells (150 mm open boreholes) drilled into the fractured bedrock (F4: 32.9 m, confined sedimentary bedrock; PO4:50 m, unconfined crystalline bedrock; PO5: 60.9 m, unconfined sedimentary bedrock), and two observation wells (2" PVC piezometers) in the Quaternary sediment (S1: 8.2 m, confined; and S5: 16.8 m, semi-confined, with screen lengths of 3 and 1.5 m respectively). Larocque et al. (2015) performed pumping tests in observation wells F4, PO4, and PO5, and slug tests in observation wells S1 and S5. They also compiled hydraulic conductivities (K) available

253 from different consulting firms (see Table 1). K values for the crystalline bedrock vary
254 between $1.0x10^{-6}$ and $8.3x10^{-5}$ m s^{-1}. These values are higher than those from the
255 literature, which can range from $8.0x10^{-9}$ to $3.0x10^{-4}$ m s^{-1} (Domenico and Schwartz
256 1990). K values for the sedimentary bedrock were highly variable spatially, and reported
257 values range between $1.1x10^{-7}$ m s^{-1}, for wells in the Covey Hill and the Potsdam
258 Formation, to $9.0x10^{-4}$ m s^{-1}, for a well in the Beekmantown Group. Reported sand K
259 values vary between $7.0x10^{-6}$ and $7.4x10^{-3}$ m s^{-1}. Measured till K values were between
260 $3.5x10^{-7}$ and $1.1x10^{-4}$ m s^{-1}. The few measured K values for clay deposits range between
261 $1.6x10^{-10}$ and $9.7x10^{-8}$ m s^{-1}.

262 Groundwater levels measured once in private wells between July and August 2013
263 are also reported in Larocque et al. (2015) for 149 locations within the Raquette River
264 watershed and 35 locations in the surrounding area (see Figure 1). Heads are also
265 available from the five monitored observation wells, and have amplitudes (maximum
266 minus minimum value during the study period) of 0.52, 6.39, and 0.80 m for bedrock
267 wells F4, PO4, and PO5 (well PO4 is located at the highest altitude on Mount Rigaud),
268 and amplitudes of 0.28 and 0.95 m for observation wells CPT1 and CPT5 in the
269 Quaternary sediments. A piezometric map of the bedrock aquifer was interpolated (co-
270 kriging with topopography data in ArcGIS; see Larocque et al. 2015 for details) using all
271 available head data from the private wells and average heads from the observation wells, in
272 addition to control points, where hydraulic connections between the surface and the
273 bedrock aquifer were clear (e.g., where the bedrock outcrops in the Raquette riverbed, at
274 ponds located on Mount Rigaud, as well as along the Outaouais River).

275 The piezometric map (Figure 2b) shows that groundwater flow is radial, with high

hydraulic gradients (0.2 m m^{-1}) on Mount Rigaud. Hydraulic gradients are small in the clay plain (0.003 m m^{-1}), where artesian conditions were observed. The piezometric map is not modified by the river channel, indicating that there is no hydraulic connection between the Raquette River and the confined bedrock aquifer in this area. North of Mount Rigaud, groundwater flows toward the Outaouais River, with a low hydraulic gradient (0.004 m m^{-1}). Groundwater flow directions between Mount Rigaud and the Saint-Lazare Hill converge towards the Raquette River (no head forcing at this location). Flow rate measurements at different locations in this area, made in August 2014, show a marked increase in river flow. Between the Saint-Lazare and Hudson hills, head control points equal to river water levels were used in the interpolation to compensate for the small number of measured heads (see Figure 1) and to reproduce the drainage effect of the river in the piezometric map. On Hudson Hill, Technorem (2005) showed the presence of two distinct aquifers, the top granular aquifer and the fractured aquifer, confined by till and clay horizons. Head values suggest the presence of a local vertical gradient between the two aquifers, but this is not confirmed at some locations.

Meteorological conditions

Long-term data for the 1981-2010 period from the Mount Rigaud weather station (MDDELCC 2014) show that total annual precipitation is 999 mm, of which 16% falls as snow between mid-November and mid-April. Total annual precipitation was 1049 mm between November 2012 and October 2013, and 1119 mm between November 2013 and October 2014. November 2012 was particularly dry, with only 7 mm of rain, while April 2014 was very wet, with 163 mm of rain. During both years

299 of the study period, June was the month with the highest precipitation, and snowmelt
300 occurred in April. Monthly average air temperature was lowest in February (-8.4°C in
301 2013 and -11.9°C in 2014), and highest in July (21°C in 2013 and 19.4°C in 2014; see
302 Figure 4). The 2013-2014 hydrological year was hotter (5.9°C) on average than the
303 2012-2013 hydrological year (4.9°C). It is important to note that, in the current work,
304 winter precipitation occurring as snow was not adjusted for wind-induced undercatch
305 and snow sublimation. These processes typically reduce the water equivalent of the
306 snow available at the time of snowmelt.

Figure 4. Net monthly vertical inflows (VI_{net}) and air temperature (Temp.) in the Raquette River watershed for the two hydrological years studied (Nov. 2012 – Oct. 2013 and Nov. 2013 – Oct. 2014). Vertical inflows (*VI*) correspond to the sum of precipitation occurring as rain when air temperature exceeds the freezing point and snowmelt occurring during the winter and in the spring snowmelt period. VI_{net} corresponds to *VI* values from which evapotranspiration has been substracted. VI_{net_cal} corresponds to the calibrated VI_{net} values.

307

308 **Methods**

309 ***Vertical inflows, and hydrological and hydrogeological monitoring***

310 In a cold region, estimating the amount of water that can potentially infiltrate or runoff is
311 crucial. Vertical inflows (*VI*) are defined as the sum of liquid precipitation and snowmelt
312 available daily for runoff, percolation, or evaporation. In the absence of snow on the ground
313 and when the precipitation is liquid, *VI* and precipitation data are equal. In the Raquette
314 River watershed, monthly *VI* was estimated from November 2012 to October 2014 (see

Figure 4) using the MOHYSE model (Fortin and Turcotte 2007), which uses daily air temperature and accumulated snow to calculate snowmelt using a degree-days approach. MOHYSE is a conceptual hydrological model that is not spatially discretized, designed to reproduce river flows at the watershed scale. After calibration on flow rates, the model provides an estimate of actual evapotranspiration (ETR), runoff, and infiltration (see Larocque et al. 2015 for details on the MOHYSE model application for the Raquette River watershed). During the study period, annual ETR was equal to 581 mm (2012-2013) and 459 mm (2013-2014). Subtracting monthly ETR from monthly *VI* values provides a VI_{net} time series, considered to be the volume of water that can potentially runoff or infiltrate in a given month. For the two years of the study period, VI_{net} values were 518 and 686 mm. The daily VI_{net} values necessary for the flow model were calculated using a different VI/VI_{net} ratio for each month.

Groundwater levels in the five observation wells that were equipped with automated *Solinst* level loggers were monitored continuously at an hourly time step by Larocque et al. (2015) from June (F4), August (PO4 and PO5), and September (S1 and S5) 2013, until October 2014 (all wells except PO5, for which monitoring ended in July 2014).

Groundwater flow model

The integrated model was developed with MikeSHE (DHI 2007). The simulated domain (see Figure 5) covers an area of 118 km^2. It is similar to that of the Raquette River watershed, but slightly different in the northeast area to allow for the representation of groundwater outflow to the Outaouais River. It is also different in the southwest portion of the study area, where the simulated domain was based on the groundwater flow divide

338 estimated from the piezometric map. The simulated domain was discretized horizontally
339 into 100 m x 100 m cells to provide a refined representation of the Raquette River. This
340 fine spatial resolution allows river entrenchment to be adequately described (Refsgaard
341 1997), but significantly increases calculation times (Vazquez et al. 2002). Surface
342 topography was described using a single LiDAR datum for each cell. Topography in the
343 model varies between 23 m close to the Outaouais River and 225 m at the top of Mount
344 Rigaud. Vertically, the numerical model simulates a porous media aquifer to a uniform
345 depth of -100 m. The model is divided into 10 layers of variable thicknesses. Layer depth
346 increases gradually, from 2.5% of the total model thickness at any given location to 20% of
347 the total model thickness. In total, 165,870 cells were used to horizontally and vertically
348 discretize the study area.

Figure 5. Boundary conditions used to represent the Raquette River study area in the integrated MikeSHE model.

349

350 ***Representing the aquifer and the surface flow network***

351 In the model, littoral sands and sand and gravel were simplified into a single sand unit.
352 The bedrock was divided into crystalline and sedimentary (essentially Potsdam
353 sandstone) bedrock, since the two formations have different hydraulic properties.
354 Because the thickness of the sandstone above the crystalline bedrock is unknown, the
355 crystalline bedrock was gradually reduced in thickness over a 2 km distance, starting
356 from the sandstone-crystalline bedrock interface. Data available to constrain the
357 conductivity calibration for the five different lithologies (i.e., sand, till, clay, crystalline
358 bedrock, and sandstone bedrock) are reported in Table 1. For each geological unit,

359 vertical hydraulic conductivity was set as 10% of the horizontal value, as is typically done
360 when no measured K_v/K_h ratio is available.

361 The Raquette River and its three main tributaries were simulated using the Mike11
362 model. A cross-section was used at each significant change in bed slope to represent
363 changes in river flow dynamics. Given the lack of precise field data, aquifer-river
364 exchanges were set as occurring at the river bed. Two parameters characterize flow in
365 the river channel, the Manning roughness coefficient (n) and the river leakage
366 coefficient (r_l). Although n affects water level in the river, and thus aquifer-river
367 interactions, the exchange coefficient has a larger influence on the exchanged fluxes
368 (Refsgaard 1997). The n for the Raquette River and its main tributaries was set as
369 20 $m^{1/3}$ s^{-1}, and their exchange coefficient was a calibrated parameter. The small
370 streams were represented using fixed head drains (see Figure 5) to evacuate water
371 when heads are above the drain elevation. The elevation of these draining cells was
372 determined using the LiDAR data. The drainage constant (C_d), which is known to
373 have a considerable effect on the form of the hydrograph (Vazquez et al. 2002), was a
374 calibrated parameter. DHI (2007) suggest C_d values of between $1.0x10^{-7}$ and $1.0x10^{-6}$ s^{-1}, although values as low as $2.0x10^{-8}$ s^{-1} (Al-Khudhairy et al. 1999) and as high as $4.9x\ 0^{-4}$ s^{-1} (Zhou et al. 2013) have been used elsewhere.

377 The detention storage was set to the MikeSHE default value of 25.4 mm (Frana 2012)
378 to represent the water depth above which runoff occurs. Manning roughness coefficients
379 (n) associated with surface runoff were attributed spatially as a function of the
380 geological unit and land use in the top numerical layer. On the Raquette River, till, sand,
381 and bedrock zones are generally associated with wooded areas, while clay is associated

with agricultural areas. The Manning roughness coefficient was fixed at 1.7 $m^{1/3}.s^{-1}$ for wooded areas (McCuen 2004). The agricultural areas are characterized by a complex drainage network of tile drainage (drains and agricultural ditches), which plays an important role in the watershed hydrology and rapidly evacuates water from the agricultural areas towards the Raquette River, shortening the hydrograph curve (Feyen et al. 2000). The Manning roughness coefficient influences river flow rates similarly to a tile drainage network.

Boundary conditions

The simulated domain does not correspond exactly to the Raquette River watershed (Figure 5). Rather, the boundary conditions were set to correspond to groundwater flow conditions. In the clay plain south of Mount Rigaud, a Neuman-type boundary condition was set at a right angle to groundwater flow, with a hydraulic gradient of 0.003 $m.m^{-1}$. To the north of Mount Rigaud, the Outaouais River was used as a Dirichlet-type constant head condition. Water level data from gauging stations located below the Carillon dam (Hydro-Québec 2014) and at the Deux-Montagne Lake (HYDAT 2014) were used to set a time-varying head. These time-varying head conditions were attributed to the top four numerical layers, equivalent to approximately 25 m below the surface, because the Outaouais River reaches approximately this depth at its center (Québec-Pêche 2011). A no-flow boundary was attributed to the lowest six layers. The other limits were flow boundaries, coinciding with either the watershed boundary (no-flow) or with boundaries imposed by groundwater flow directions (imposed flux).

Unsaturated flow and evapotranspiration were not simulated in the model so as to

reduce the number of unknown parameters and the simulation time. Alternatively, a simple algebraic representation of recharge was used (adapted from DHI 2007):

$$Rec = VI_{net} f_I \quad \text{eq. (1)}$$

where Rec is recharge (mm d^{-1}), VI_{net} is the water available for runoff or infiltration (mm d^{-1}), and f_I is the fraction of available water that can infiltrate (unitless). f_I is constant in time but varies spatially with the surface geology (bedrock, till, clay, and sand).

Simulated conditions

The transient-state simulation spanned two hydrological years, from November 1st 2012 to October 31st 2014. A ten-year spin-up period, in which the meteorological data for the two years of the study period were repeated five times, was used to ensure that initial conditions were stable prior to simulating the actual conditions. Simulated flow rates and heads for over these two years were compared to measured values, and the model was calibrated manually. The time step was one day.

The model was calibrated through manual trial-and-error using the available head measurements for the modeled domain and using the hourly groundwater levels that were measured at the five observation wells. The flow rate data are from the three gauging stations. All these data were reported in Larocque et al. (2015). The relatively large number of calibrated parameters (22; Table 1) makes it plausible that the combination of parameters that satisfy the minimization criteria could be numerous. Using an automatic

calibration procedure might have alleviated this situation, but would not have eliminated it. The calibrated parameters are thus considered to represent one set of possible values.

The model was calibrated using a total of 22 parameters: five hydraulic conductivities (K) and five storage coefficients (S_s and S_y) of the different geological units (sand, till, clay, sandstone, and crystalline bedrock), the river leakage coefficient (r_l), the drainage constant for the small streams (C_d), the Manning roughness coefficient for runoff (n) (clay), and the infiltration fractions (f_l) for the sand, till, clay, and bedrock outcrops. Calibration intervals are described in Table 1, and are either measured values for the study area or from the scientific literature. The calibration of hydraulic heads was done to minimize the mean error (ME), the mean absolute error (MAE), the root mean square error (RMSE), and the normalized root mean square error (NRMSE).

$$ME = \frac{1}{N}\sum_{i=1}^{N}\left(h_{s_i} - h_{m_i}\right) \quad \text{eq. (2)}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|h_{s_i} - h_{m_i}\right| \quad \text{eq. (3)}$$

$$RMSE = \frac{1}{N}\sqrt{\sum_{i=1}^{N}\left(h_{s_i} - h_{m_i}\right)^2} \quad \text{eq. (4)}$$

$$NRMSE = \frac{RMSE}{(h_{max} - h_{min})} \quad \text{eq. (5)}$$

where N is the number of head measurements, h_{si} is the simulated head (m), h_{mi} is the measured head (m), h_{max} is the highest measured head (m), and h_{min} is the lowest measured head (m).

The model performance for surface flow was evaluated using the following criteria:

$$NSE = \frac{(Q_m - \bar{Q}_m)^2 - (Q_s - \bar{Q}_s)^2}{(Q_m - \bar{Q}_m)^2} \quad \text{eq. (6)}$$

$$NSE_{log} = \frac{\left(log(Q_m) - log(\bar{Q}_m)\right)^2 - \left(log(Q_s) - log(\bar{Q}_s)\right)^2}{\left(log(Q_m) - log(\bar{Q}_m)\right)^2} \quad \text{eq. (7)}$$

$$F_{Bal} = \frac{\bar{Q}_m - \bar{Q}_s}{\bar{Q}_m} \quad \text{eq. (8)}$$

$$PEP = \sum_{i=1}^{5} \frac{Q_{max\_m(i)} - Q_{max\_s(i)}}{Q_{max\_m(i)}} \quad \text{eq. (9)}$$

Where N_Q is the number of measured flow rates, Q_m is the measured flow rate ($m^3\ s^{-1}$), Q_s is the simulated flow rate ($m^3\ s^{-1}$), Q_{m_bar} is the average measured flow rate ($m^3\ s^{-1}$), and Q_{s_bar} is the average simulated flow rate ($m^3\ s^{-1}$).

The general quality of the simulated flow rates was estimated using the Nash-Sutcliffe model Efficiency coefficient (*NSE*; Nash and Sutcliffe 1970). Because the *NSE* is particularly sensitive to high flow rates (Güntner et al. 1999), the *NSE* calculated on log-transformed flow rates (*NSElog*) was used to evaluate model performance during low flow periods (Güntner et al. 1999). The ability of the model to simulate the average and maximum flood flow rates at each gauging station was estimated using the *Fbal* criteria (Henriksen et al. 2003) and the Percent Error in Peak (*PEP*; Green and Stephenson 1986) respectively. To evaluate the general ability of the model to simulate flood events, the PEP criteria were averaged, with the *PEP* calculated for the five largest flow rates measured. Ideally, *Fbal* and *PEP* should tend toward zero, while *NSE*, *NSElog*, and *NSEcomb* should tend toward one. *Fbal* is positive when a simulated flow rate is smaller

than the observed value, and vice versa. Similarly, a negative *PEP* indicates that the simulated flow rate is larger than the measured flow rate, and vice-versa.

The model sensitivity coefficients were calculated for each parameter using a perturbation of 5% applied to the calibrated parameter values listed in Table 1. To facilitate the process, the four f_1 values (infiltration fraction) were modified simultaneously, resulting in a total of 19 tested parameters. The calculations were done using the AUTOCAL module in Mike SHE:

$$S_{i,j} = \frac{F(\theta_1,\theta_2,\ldots,\theta_i+\Delta\theta_i,\theta_n)-F(\theta_1,\theta_2,\ldots,\theta_i-\Delta\theta_i,\theta_n)}{2\Delta\theta_i} \quad \text{eq. (10)}$$

$$S_{i\_TOT} = \sum_{j=1}^{k} S_{i,j} \quad \text{eq. (11)}$$

Where $S_{i,j}$ is the sensitivity of model result j to parameter θ_i. Subscript j varies from 1 to 3 (k) in this case for flow rates (three flow rate measuring stations), and from 1 to 5 (k) for heads (five observation wells). The variable n corresponds to the total number of tested parameters (n=19), $\Delta\theta_i$ is the 5% perturbation imposed on a parameter θ_i, and S_{i_TOT} is the total sensitivity of the model to this parameter. The sensitivity calculations were done for flow rates and heads separately, leading to S values for each of these two types of results. Prior to the sensitivity calculations, logarithmic transformations were applied to parameters with ranges spanning two or more orders of magnitude (i.e., K, S_s, n, and r_l).

Results

Calibrated parameters

The hydraulic conductivities calibrated for sand, till, and sedimentary bedrock are within

484 the ranges of measured values (Table 1). Calibrated clay *K* was slightly higher than the
485 measured range, while calibrated crystalline bedrock *K* is one order of magnitude lower
486 than the measured range. The hydraulic conductivity of the sedimentary bedrock is
487 significantly higher than that of the crystalline bedrock, as expected for this more
488 productive formation.

489 Storage coefficients and specific storage values for all of the geological units are
490 within the range of available literature values, with the exception of the sedimentary
491 bedrock *S* and the sand S_s. Although there are unconfined areas in the study area, storage
492 coefficients represent average values over the watershed, since geological units have
493 been greatly simplified.

494 Although the river-aquifer exchange coefficient is known to vary spatially and
495 temporally due to erosion and sedimentation processes, as well as water temperature and
496 water levels in the river (Doppler et al. 2007), a single value of 7.0×10^{-7} s^{-1} was
497 calibrated for the entire length of the river. The drainage constant affects the high flows
498 and controls the speed with which the small streams drain groundwater (Sahoo et al.
499 2006). In the study area, the calibrated drainage constant was 5.0×10^{-6} s^{-1}, within the
500 range of measured values.

501 For sand, till, and bedrock units associated with wooded areas, the Manning
502 roughness coefficient (*n*) was not calibrated. It was calibrated for clay areas to
503 reproduce the maximum flow rates encountered during flood events. A high value was
504 necessary to evacuate water rapidly through runoff from clay-covered areas. It is
505 hypothesized that this rapid mobilization of water at the surface has a similar impact on
506 river flow rates as that of the agricultural tile drainage system, which was not otherwise

represented in the model. The calibrated value of 15 $m^{1/3}$ s^{-1} appears to be realistic, since it is higher than that used for the wooded areas (1.7 $m^{1/3}$ s^{-1}), and lower than that proposed for river flow (20.0 $m^{1/3}$ s^{-1}).

The infiltration fractions were also distributed spatially as a function of the geology of the top numerical layer. The values were calibrated to best represent the simulated heads. The infiltration fractions calibrated for the sand, till, clay, and undifferentiated bedrock were 0.70, 0.42, 0.05, and 0.60 (Table 1). These values are comparable to those obtained by Croteau et al. (2010) over a 39-year simulation of the Chateauguay River watershed. The infiltration fraction obtained here for the undifferentiated bedrock is slightly lower than that of Croteau et al. (2010).

The calibrated model was fine-tuned by modifying the monthly VI_{net} data to provide a better adjustment for spring flow rates. This was done by reducing the readily available rain that occurred during the winter months (November to March) by half. The other half was considered to be stored in the snowpack, to the made available during spring snowmelt (April). This proportion of winter-available precipitation that was transferred instead to water available in the spring resulted from a manual trial-and-error process. This volume represents water that falls on snow and freezes without reaching the river. Considering that there is limited recharge over frozen soil, the snowmelt associated with winter thawing is thus redistributed to the spring flood.

The flow rates show limited sensitivity (<1 and > 0.1) to sand K (S_{TOT} = -0.66), crystalline bedrock K (S_{TOT} = -0.49), sedimentary bedrock S_y (S_{TOT} = -0.32), clay K (S_{TOT} = 0.26), sedimentary bedrock K (S_{TOT} = 0.14), till K (S_{TOT} = -0.14), the C_d coefficient (S_{TOT} = 0.13), the r_l coefficient (S_{TOT} = 0.10), and the n coefficient (S_{TOT} = -0.10). The

model shows almost no sensitivity to the other 10 parameters ($S_{TOT} < 0.1$ in all cases; Figure 6a). The heads are most sensitive to crystalline bedrock K ($S_{TOT} = 73$) and sedimentary bedrock K ($S_{TOT} = -23$), and, to a more limited extent, to sand K ($S_{TOT} = -6.5$) and clay K ($S_{TOT} = -1.3$). The heads show limited sensitivity (<1 and > 0.1) to crystalline bedrock S_s ($S_{TOT} = -0.85$), clay S_y ($S_{TOT} = 0.41$), the n coefficient ($S_{TOT} = 0.18$), till K ($S_{TOT} = -0.17$), and sand S_y ($S_{TOT} = -0.11$). The model shows almost no sensitivity to the other 10 parameters ($S_{TOT} < 0.1$; Figure 6b). The apparently higher total sensitivity of heads to parameter variation, compared with the less sensitive flow rates, could be related to the fact that S_{TOT} is the sum of S_i values for five observation wells, whereas it is the sum of S_i values for only three flow rate measuring stations.

Positive total sensitivity values indicate a dominance of positive sensitivities (i.e., the head or flow rate increases when a parameter increases), whereas negative total sensitivity values indicate a dominance of negative sensitivities (i.e., the resulting head or flow rate increases when a parameter decreases). Interestingly, the total sensitivity can be positive for flow rates while being negative for heads (e.g., in the case of clay K). Again, this could be related to the fact that there are less flow rate measuring stations than observation wells.

Figure 6. Absolute values of total sensitivity coefficients (S_{i,j_TOT}) with respect to a) flow rates and b) heads, corresponding to a 5% parameter variation for a) flow rates, and b) heads. Plus signs indicate a positive sensitivity, whereas no sign indicates a negative sensitivity.

549 ***Simulated flow rates***

550 The integrated model simulated measured flow rates at the three gauging stations relatively
551 well, with NSE values of 0.72, 0.75, and 0.62 for stations 1, 2, and 3 respectively
552 (Figure 7). The NSElog values of 0.65, 0.53, and 0.12 indicate that the low-flows are well-
553 simulated at stations 1 and 2, but that they are less well reproduced at station 3. The
554 minimum simulated flow rates at stations 1, 2, and 3 are 0.34, 0.21, and 0.04 $m^3.s^{-1}$
555 respectively, compared to measured low flows of 0.40, 0.10, and 0.03 $m^3.s^{-1}$. Although the
556 minimum flow rate is underestimated at station 2, these simulated flow rates are of the
557 correct order of magnitude.

Figure 7. Measured and simulated flow rates at gauging stations a) 1, b) 2, and c) 3.

558

559 Positive Fbal values, 0.34, 0.07, and 0.21, for gauging stations 1, 2, and 3 respectively,
560 indicate that the average flow rates are underestimated by the model at all three stations.
561 The larger Fbal at station 1 is due to the significant underestimation of the spring flow rates
562 in 2013 and 2014, which were measured only at this station. The PEP calibration criteria,
563 which evaluates the quality of the flood simulation, were 0.42, 0.33, and 0.55 for stations 1,
564 2, and 3 respectively. This indicates that, on average, the five highest flow rates measured
565 were underestimated at the three stations, and particularly at station 3.

566 During summer months when VI_{net} values are not equal to zero (i.e., $VI \leq$ ETR), the
567 model does not simulate small flow rate variations, such as those observed at the end of
568 July 2013. The model also overestimates the flow observed in September 2013, while
569 flows increase markedly between March and July of both years, as well as between
570 September and November 2013, and in October 2014.

Flow rates at the outlets of the three tributaries were simulated in the same way as those in the Raquette River (Figure 8). They were found to vary significantly over the study period, between 0.002 and 6.46 $m^3 s^{-1}$ for tributary 1, between 0.001 and 13.70 $m^3 s^{-1}$ for tributary 2, and between 0.001 and 4.48 $m^3 s^{-1}$ for tributary 3. Tributaries 1 and 3 have similar flow rates, while the peak flow rates are the higher for tributary 2. All three tributaries have very low flows in July and August 2013 and 2014. The values simulated for August 2014 are of the same order of magnitude as the corresponding measured values: an average simulated flow rate of 0.002 $m^3 s^{-1}$ compared to a measured value of 0.008 $m^3 s^{-1}$ for tributary 1, an average simulated flow rate of 0.004 $m^3 s^{-1}$ compared to measured values of 0.002 and 0.003 $m^3 s^{-1}$ for tributary 2, and an average simulated flow rate of 0.001 $m^3 s^{-1}$ compared to measured values of 0.002 and 0.004 $m^3 s^{-1}$ for tributary 3. No flow rates were measured at the tributaries' outlets at other times.

Figure 8. Simulated flow rates at the outlets of tributaries 1, 2, and 3.

Simulated heads

The average heads simulated with the calibrated model for the two years were found to adequately represent the 149 one-time measured heads in the Raquette watershed, and the average heads measured in the five observation wells (Figure 9). The mean error (ME) of -1.24 m indicates no systematic over- or underestimation of heads, and the MAE of 6.85 m is of the same order of magnitude as the maximum head amplitude in the monitored observation wells (6.39 m at the crystalline bedrock well, PO4, on Mount Rigaud). The RMSE is 8.77 m and the NRMSE is 0.06, a value smaller than the 0.1 target value (Gallardo et al. 2005; Lutz et al. 2007). However, the largest simulated

errors were not located in the crystalline bedrock of Mount Rigaud, but were found for Hudson Hill (-29.54 m) and for the eastern and northern extents of Mount Rigaud (-26.70 and -19.10 m respectively). Some heads simulated for the plain north of Mount Rigaud, close to the Outaouais River boundary, are overestimated by approximately 12 m.

Figure 9. Measured and simulated steady-state heads. The thick black line is the 1:1 line and the dotted lines represent a +/- 5 m error envelope.

The transient heads simulated for the five observation wells were plotted relative to surface elevation to remove any errors related to topographical inaccuracies (Figure 10). Head variations (timing and amplitude) at wells S1 and PO5 (unconfined conditions in sand deposits and in the sedimentary bedrock) were correctly simulated. The transient-state head simulations were of lower quality for wells S5 and PO4 (semi-confined conditions in the clay plain and unconfined condition in the crystalline bedrock), and were even poorer for well F4 (confined conditions in the clay plain). The low VI_{net} values used to simulate the winter seasons, followed by the large VI_{net} values at snowmelt, have a visible effect on the heads simulated at the five observation wells. This effect is apparent from the measured head time series at two unconfined wells, S1 and PO5. The F4 well tapping the confined aquifer in the clay plain south of Mount Rigaud does not exhibit either a winter decrease or a spring increase in heads. The semi-confined S5 well shows a winter decrease and a spring increase, which are well-simulated in terms of timing, but overestimated in terms of amplitude. The simulated heads at the PO4 well in the crystalline bedrock of Mount Rigaud decrease slowly over the winter, compared with

614 almost no measured head decrease, and increase more sharply at snowmelt than was
615 measured in reality.

Figure 10. Measured and simulated transient-state heads at the five observation wells. Note that the vertical scale of the PO4 crystalline bedrock well is larger than that of the other wells.

616

617 ***Simulated groundwater discharge***

618 The draining cells representing the streams in the model (see Figure 4 for locations)
619 respond to the VI_{net} time series with maximal groundwater discharge during the spring
620 melt and gradually decreasing discharge for the months with limited or no VI_{net},
621 corresponding to the gradually decreasing groundwater levels (Figure 11a). The streams
622 located in the lower portion of the study area (between gauging stations 1 and 2) have the
623 highest flow rates, with monthly values ranging between 0.052 $m^3 s^{-1}$ (March 2014) and
624 0.136 $m^3 s^{-1}$ (April 2014). Drained flow rates in the middle portion (between gauging
625 stations 2 and 3) are much lower, ranging from 0.010 to 0.030 $m^3 s^{-1}$. Above gauging
626 station 3, the drained flow rates are very low throughout the year, except between
627 snowmelt and early summer, when they vary between 0.001 $m^3 s^{-1}$ and 0.015 $m^3 s^{-1}$ in
628 2014. The very low drained flow rates in the higher portion of the study area reproduce
629 the intermittent flow conditions observed in the small streams. Elsewhere in the study
630 area, the draining cells remain active throughout the simulation period, indicating that the
631 water table is generally higher than the drains.

Figure 11. Monthly simulated net vertical inflows (VI_{net}) and flows between gauging

stations 1 and 2, between gauging stations 2 and 3, and above gauging station 3 for a) drained flows and b) baseflows.

632

633 The groundwater flow contribution to the Raquette River reacted to VI_{net} similarly to
634 the drained flows, albeit with a lower amplitude (Figure 11b). The central and lower
635 portions of the simulated area responded similarly to VI_{net}, with reductions extending from
636 mid-summer (August) to the end of winter (March) in the two simulated years. The rate of
637 base flow reduction is relatively constant through time, notwithstanding positive VI_{net}
638 months. The base flows increase in April and May, following the spring snowmelt, and
639 are constant in June and July. In the lower portion of the watershed, the aquifer
640 contributes with average monthly inflows of between 0.048 $m^3 s^{-1}$ (March 2013 and 2014)
641 and 0.070 $m^3 s^{-1}$ (July 2014). The central portion of the river, between gauging stations 2
642 and 3, is where the aquifer contributes most to the river in the Sainte-Marthe Valley,
643 with average monthly groundwater inflows of between 0.086 and 0.110 $m^3 s^{-1}$. The
644 aquifer contribution to the river upstream of gauging station 3 is relatively constant
645 throughout the simulated period, on average 0.015 $m^3 s^{-1}$.

646

647 **Discussion**

648 ***Model sensitivity***

649 This work has shown that, for both flow rates and heads (see Figure 6), the model is less
650 sensitive to S_s and S_y for sand, till, clay, and bedrock than it is to K values for the same
651 geological formations. This suggests that data acquisition aimed at developing an integrated
652 flow model should prioritize the quantification of hydraulic conductivities, through

pumping tests, for example. The following calibration could be divided into a sequential process, starting with the calibration of K values to reproduce flow rates and heads, followed by a subsequent fine-tuning with calibration of S_s, S_y, r_l, C_d, and n coefficients. It is important to highlight that the latter three parameters are rarely measured in the field, and are instead usually estimated from the scientific literature.

Interestingly, the sensitivity analysis showed that river flow rates are influenced by both hydraulic properties (K and S_y had $S_{TOT} > 0.1$) and surface flow parameters (r_l, C_d, and n had $S_{TOT} > 0.1$). Heads are influenced mainly by hydrogeologic properties (K, S_s, and S_y had $S_{TOT} > 0.1$), but are also influenced by the surface flow parameter, n ($S_{TOT} > 0.1$). This underlines the importance of using an integrated model (coupling surface flow and groundwater flow) that considers the hydraulic properties of the surrounding aquifers and the surface flow parameters when representing the water flow dynamics over the entire watershed.

The parameters to which flow rates and heads show almost no sensitivity ($S_{TOT} < 0.1$) represent more than half of all the calibrated parameters. As could be expected, these parameters of lesser importance include S_s and S_y values for some geological formations for the flow rates, and include the r_l and C_d coefficients, as well as S_s and S_y values for some geological formations for heads. Interestingly, the f_I coefficient (infiltration fraction) had almost no effect on either flow rates ($S_{TOT} = -0.0067$) or heads ($S_{TOT} = 0.0013$). This indicates that this parameter, which should control the volume of water that infiltrates and flows through runoff, does not adequately represent the partition between these two processes. This is not to say that this partitioning is not important in integrated flow modeling, but a complete representation would probably reveal a greater impact of this

process on flow rates and heads.

It is important to note that the model calibration was performed using a trial-and-error approach. Although this method can provide a reasonably well-calibrated model, using an automatic calibration approach might have identified a different parameter combination that would have provided a similar or even improved flow rate and head calibration.

Simulated flow rates and heads

Overall, the simulated flows were lower than the measured flows. For the frost-free period, from April 1st to October 31st 2013, the simulated total flow at station 1 was 256 mm, while the measured value was 380 mm. For the same period in 2014, the simulated and measured total flows were 385 and 557 mm respectively. These discrepancies might be explained by the simplified representation of evapotranspiration and infiltration processes, which might have underestimated the available water in the model. They could also have been caused by the uncertainty of the rating curve for high flows at station 1. The high flows occur only during a short period of time in the March-to-May spring snowmelt and during large summer rain events. However, because they are much larger than the baseflows, they can result in an overestimation of the total annual flow. The maximum flow rate at station 3 (furthest up gradient) may have been higher than that at station 2 (middle) due to errors in the rating curves at these two gauging stations. Performing a simulation over a longer time period would also allow the model to be verified under a wider range of hydrological conditions.

Part of these simulation errors can be attributed to the simplified representation of VI_{net}. Because evapotranspiration is not simulated explicitly, the available water every month has

699 been reduced using an evapotranspiration fraction. This results in large and intense rain
700 vents being distributed uniformly over a month, which causes an underestimation of runoff
701 and flow rate during and after precipitation events. This problem was noted by Jones et al.
702 (2008), who reduced precipitation using a constant evapotranspiration fraction. In reality,
703 during rain events, the evapotranspiration rate is one or many orders of magnitude smaller
704 than the precipitation rate (Camporese et al. 2010). Moreover, the spatial variability of
705 interception and evapotranspiration processes are overly simplified in the Raquette River
706 model, and are not a function of land use cover. Using the VI_{net} artifact to estimate
707 available water for runoff and infiltration thus appears to be sound regionally, but is not
708 entirely satisfactory. However, the alternative of using detailed information to represent
709 different land use covers would have added an unwarranted level of complexity to the
710 model.

711 According to Furman (2008), using an infiltration fraction that is constant in time is
712 less precise than explicitly simulating infiltration in the unsaturated zone, which produces a
713 delay between infiltration and actual groundwater recharge, and considers soil water content
714 in estimating the volume of water to runoff and infiltrate for each rain event. The simple
715 approach used in this work instantaneously transfers a fraction of the rain to the saturated
716 aquifer, not considering soil water content. During intensive rain events, this approach can
717 significantly reduce the amount of runoff, which can lead to an underestimation of
718 simulated peak flow rates in the river. Again, a complete representation of the infiltration
719 processes might prove to more effectively fine-tune river flow rates and heads.

720 In the model, accelerated runoff with a low roughness coefficient has been used to
721 mimic the role of agricultural drains (not represented explicitly in the model) that

722 quickly discharge water following rain events. However, natural runoff is influenced by
723 the topography, and is therefore not directly routed to streams, as would be the case for
724 agricultural drains that are usually connected directly to streams and rivers. This may be
725 another explanation for the underestimation of simulated flows, particularly in the clay
726 plain upstream of the study area, where the slopes are very low.

727 The differences between measured and simulated flows could also be explained in
728 part by differences between the estimated watershed boundary and the position of the
729 model boundary (see Figures 1 and 4), as the actual catchment at station 1 covers an
730 area of 133 km^2, while the simulated basin covers 118 km^2. However, this is expected to
731 have a more limited effect on simulated flows than the above-mentioned causes. Longer
732 time series over many years would provide further insight into what causes the
733 differences between simulated and measured flow rates.

734 The adequacy of simulated flow rates at the outlets of tributaries 1, 2, and 3 could
735 not be verified, because measurements were only available for the low flow conditions
736 encountered in August 2014. These low flow conditions appear to be reasonable, and
737 there is no indication that peak flows are over- or underestimated. In the absence of
738 more measured flow rates, using a different numerical representation for these three
739 tributaries (rivers with exchanged flows) compared to other, smaller streams (drains)
740 might not be justified.

741 The simulated errors on the average simulated heads associated for the plain north of
742 Mound Rigaud and close to the Outaouais River could be related to the use of a no-flow
743 condition for the six lower layers of the aquifer. The impact of boundary conditions on
744 simulated heads is a commonly reported problem (e.g., Sulis et al. 2011). For the

745 Raquette River, this overestimation of heads in this part of the aquifer might have
746 induced simulated baseflows that are higher than actual values, but this was not
747 evidenced with available flow rate measurements. In the Raquette River model, tests
748 performed with a constant head boundary condition on all 10 layers resulted in a lower
749 error on heads in the vicinity of the Outaouais River. However, because there is no
750 indication that groundwater enters the river below a 25 m depth, using a no-flow lower
751 boundary appeared to be more realistic. In other areas, the overall error on the average
752 simulated heads, especially for those above 75 m, which are exclusively from the
753 crystalline bedrock, indicates that the spatially homogeneous hydraulic properties used in
754 the porous equivalent model represent average values that do not reflect locally higher or
755 lower values, due to the presence of flow in the fractured media.

756 The transient heads simulated at well F4, and to a lesser extent at well S5, decrease
757 more rapidly and markedly than heads measured during the 2013-2014 winter months
758 (November to March), indicating that the simulated aquifer releases too much water in
759 this area during the snow covered period. The overly simplified spatial distribution of
760 storage coefficients may be the cause of this discrepancy, but this could not be
761 overcome in the model without over-parameterizing, due to the availability of only five
762 continuously measured piezometers. The simulated and measured heads at the PO4
763 crystalline bedrock well on Mount Rigaud vary much more in amplitude than those of
764 the PO5 sedimentary bedrock well, located on the Saint-Lazare Hill. This is probably
765 due to the fact that the storage coefficient of the crystalline bedrock (PO4 well) is
766 smaller than that of the sedimentary bedrock (PO5 well). The use of a porous medium
767 to represent the fractured (crystalline or sedimentary) bedrock aquifer could also

768 explain part of the difference between measured and simulated transient-state heads.
769 The complex geological setting of the study area, including low-K crystalline bedrock,
770 highly transmissive sedimentary bedrock, high-K sand deposits, and substantial clay
771 deposits contribute to the challenge of simulating transient-state heads.
772
773 ***Groundwater flow contribution to the Raquette River***
774 Because they were obtained after a few days without precipitation events, the flow rates
775 measured during the August 2014 low flow conditions provide relevant information about
776 the groundwater flow contribution to the Raquette River. The sharp increase in flow rates
777 after station 8 could be due to a contribution from the aquifer through relatively
778 permeable and heterogeneous materials within the clayey river bed deposits. That this
779 inflow was contributed by the aquifer instead of drain flows was confirmed by the high
780 ^{222}Rn activities in surface water found in this portion of the river (Moreira 2016). In the
781 Sainte-Marthe Valley, between stations 9 and 19, the relatively constant rate of flow
782 increase indicates a similar level of connectivity between the river and the sand and
783 bedrock aquifer, which outcrops along the reach where the river has eroded the clay
784 deposits. The decrease in river flow between stations 19 and 20, and between stations 21
785 and 23, indicates that the river discharges water to the fluvio-glacial sand deposits in these
786 areas.
787 The relatively limited variation in baseflow contributions to the river through time
788 indicates that the simulated groundwater fluxes to the Raquette River are highly dependent
789 on storage in the saturated aquifer which apparently contributes a relatively constant
790 groundwater flux to the river through the seasons. The absence of baseflow variation

upgradient from station 3 reflects the limited aquifer-connectivity in this area.

The simulated baseflows to the Raquette River are greatest between stations 2 and 3 (Figure 11b), corresponds with measured values. In August 2014, the measured baseflows in this reach varied between 0.03 and 0.15 $m^3 s^{-1}$, while simulated values in the area are 0.10 $m^3 s^{-1}$. Between stations 1 and 2, the simulated baseflows are 0.065 $m^3 s^{-1}$, while measured values range between 0.15 and 0.21 $m^3 s^{-1}$. This difference could be explained by the inflow of water from six streams flowing from Mount Rigaud to the Raquette River. Although their flow rates were not measured in the field, it is in this reach that the streams have the largest simulated contribution to the river (Figure 11a). In particular, this contribution could explain the increase in the measured river baseflows between stations 20 and 22, after the river has fed the aquifer over 1-2 km. Between stations 2 and 3, the model simulates very limited groundwater inflow to the river, which corresponds well with the measured values of August 2014. In light of these results, the spatial distribution of simulated groundwater contribution to the river appears reasonable along the Raquette River. The river leakage coefficient and the drainage constant can thus be considered well-calibrated in their representation of measured values. However, it is possible that spatially variable river leakage coefficients and drainage constants would have provided more representative results. These parameters are known to vary in time and space following erosion and sedimentation processes, as well as with water temperature and levels (Blaschke et al. 2003; Doppler et al. 2007). In this study, not enough calibration data were available to make this fine-tuning relevant. At all three gauging stations, the model shows only limited changes in baseflows throughout the year (no variations at gauging station 3), but this could not be compared to measured baseflow values. Additional flow rate

814 measurements during low flow at all streams would be useful to further validate the model.

815

816 ***Recommendations***

817 This work has shown that some hydraulic and surface flow parameters have an impact on
818 both simulated flow rates and simulated heads, highlighting the importance of using
819 integrated models to study water flow at the watershed-scale and to address issues related to
820 integrated water management. Although this interconnected influence of processes on flows
821 may not be universal, it is particularly important when a connectivity between rivers and
822 unconsolidated aquifers exists in post-glacial environments, such as those found in the
823 St. Lawrence Lowlands, elsewhere in southern Canada, and at similar latitudes around the
824 world.

825 This work has also emphasized the importance of including a representation of
826 unsaturated zone infiltration processes. This representation is available in models such as
827 MikeSHE, but requires more parameterization efforts in order to describe soil and plant
828 parameters, and evaporation processes. The results from this study point to the fact that
829 these additional parameters are necessary to provide the fine-tuning needed to better
830 represent the watershed reactions to rain events and snowmelt.

831 Including artificial drainage might also improve flow rate simulations. Tile drainage is
832 widely used in the St. Lawrence Lowlands, even where soils are permeable. These drains
833 are known to export significant volumes of water towards the streams, thus accelerating
834 water transit time within a watershed. Although this still needs to be demonstrated more
835 generally and beyond the local scale, tile drainage could also reduce groundwater recharge
836 under given conditions at the regional scale. However, a rigorous simulation of artificial

837 drainage would require additional exported drain flux measurements.

838 The relatively good simulation results indicate that the model is useful in is current state,
839 even though uncertainties and limitations have been identified. However, if water managers
840 are to use the model to estimate the impacts of land use and climate changes, it needs to be
841 calibrated and validated over a wider variety of hydrological and meteorological conditions.
842 In this study, the model was calibrated over two relatively similar hydrological years. It is
843 therefore recommended to revisit model calibration in a few years to refine its
844 parameterization over extended conditions, making it more robust for management
845 applications. Further model developments could also be considered, for example the
846 simulation of nutrient transport through runoff, infiltration, and tile drainage.

847

848 **Conclusion**

849 The objective of this work was to examine the challenges encountered when simulating
850 surface water – groundwater interactions in a post-glacial environment, in order to guide
851 their wider use for integrated water management. An available database was used to build
852 an integrated 3D flow model for the Raquette River drainage area in the Vaudreuil-
853 Soulanges region of southern Quebec (Canada). Simulated flows were compared to
854 measured flow rates and heads from two hydrological years.

855 This work has shown that an integrated model based on data that are relatively easily
856 available for southern Québec aquifers can provide reasonable estimates of current
857 conditions. Some discrepancies were observed between observations and simulated
858 results, with simulated total flows being lower than measured values, average heads
859 having non-negligible errors in some areas, and simulated transient-state heads with

amplitudes different from measured values at some wells. Possible improvements to the model include basing the model calibration on a longer period, including long-term calibrated rating curves and reducing the uncertainty in the rating curves, and including flow measurements at of the tributaries from throughout the year. Using a complete representation of unsaturated zone processes and including the simulation of the agricultural tile drainage system could also contribute to reducing simulation errors.

The sensitivity analysis has underlined the importance of using an integrated model to simulate groundwater – surface water exchanges for integrated water management. Using the Raquette River integrated model to its full potential will require further development and long-term maintenance which will necessitate new data. This would provide watershed managers with a robust model to assess the long-term implications of water use and land use planning under changing climate conditions.

Acknowledgements

The authors would like to thank the Quebec Ministère du Développement durable, de l'Environnement et de la Lutte contre les changements climatiques (Quebec Ministry of Environment), the Vaudreuil-Soulanges regional county municipality, and the COBAVER-VS watershed agency, who contributed funding to this research. The authors also thank the private owners who allowed access to their land for water monitoring and water sampling.

881 **References**

882 Al-Khudhairy, D.H.A., J.R. Thompson, H. Gavin, and N.A.S. Hamm. 1999.
883 Hydrological modelling of a drained grazing marsh under agricultural land use and
884 the simulation of restoration management scenarios. *Hydrological Sciences Journal*
885 44(6): 943-971.

886 Blaschke, A.P., K.H. Steiner, R. Schmalfuss, D. Gutknecht, and D. Sengschmitt.
887 2003. Clogging Processes in Hyporheic Interstices of an Impounded River, the
888 Danube at Vienna, Austria. *International Review of Hydrobiology* 88(3-4): 397-
889 413.

890 Brunke, M., and T. Gonser. 1997. The ecological significance of exchange processes
891 between rivers and groundwater. *Freshwater biology* 37(1): 1-33.

892 Camporese, M., C. Paniconi, M. Putti, and S. Orlandini. 2010. Surface-subsurface
893 flow modeling with path-based runoff routing, boundary condition-based coupling,
894 and assimilation of multisource observation data. *Water Resources Research*
895 W02512, doi:10.1029/2008WR007536.

896 Carter, R.W., and I.E. Anderson. 1963. Accuracy of current meter measurements. *J.*
897 *Hydraulics Division, Proc. ASCE* 4(1): 105-115.

898 Chen, X., and L.Shu. 2002. Stream-aquifer interactions: Evaluation of depletion
899 volume and residual effects from ground water pumping. *Groundwater* 40(3): 284-
900 290.

901 Croteau, A., M. Nastev, and R. Lefebvre. 2010. Groundwater Recharge Assessment in
902 the Chateauguay River Watershed. *Canadian Water Resources Journal* 35(4):
903 451-468.

DHI. 2007. *Mike SHE user manual, Volume 1: User guide*. Hørsholm, Danemark. 396 p.

Domenico, P., and M. Mifflin. 1965. Water from low-permeability sediments and land subsidence. *Water Resources Research* 1(4): 563-576.

Domenico, P.A., and F.W. Schwartz. 1990. *Physical and chemical hydrogeology*. New York, USA: John Wiley & Sons. 824 p.

Doppler, T., H.J.H. Franssen, H.P. Kaiser, U. Kuhlman, and F. Stauffer. 2007. Field evidence of a dynamic leakage coefficient for modelling river–aquifer interactions. *Journal of Hydrology* 347(1–2): 177-187.

Feyen, L., R. Vázquez, K. Christiaens, O. Sels, and J. Feyen. 2000. Application of a distributed physically-based hydrological model to a medium size catchment. *Hydrology and Earth System Sciences Discussions* 4(1): 47-63.

Fleckenstein, J.H., S. Krause, D.M. Hannah, and F. Boano. 2010. Groundwater-surface water interactions: New methods and models to improve understanding of processes and dynamics. *Advances in Water Resources* 33(11): 1291-1295.

Fortin V. and R. Turcotte. 2007. *Le modèle hydrologique MOHYSE*. Research report, Environnement Canada. 14 p.

Frana, A.S. 2012. *Applicability of Mike SHE to simulate hydrology in heavily tile drained agricultural land and effects of drainage characteristics on hydrology*. MSc Iowa State University, Ames, Iowa, USA. Retrieved from http://lib.dr.iastate.edu/etd/12859 no. 12859. 150 p.

Furman, A. 2008. Modeling coupled surface–subsurface flow processes: A review.

926 *Vadose Zone Journal* 7(2): 741-756.

927 Gallardo, A., W. Reyes-Borja, and N. Tase. 2005. Flow and patterns of nitrate
928 pollution in groundwater: a case study of an agricultural area in Tsukuba City, Japan.
929 *Environmental Geology* 48(7): 908-919

930 GéoMont. 2011. *Acquisition des données LiDAR - Secteurs de la Vallée-du-Haut-*
931 *Saint-Laurent et Châteauguay & Vallée-du-Richelieu.* Saint-Hyacinthe, Canada.
932 38 p.

933 Gleeson, T. and A.H. Manning. 2008. Regional groundwater flow in mountainous
934 terrain: Three-dimensional simulations of topographic and hydrogeologic controls.
935 *Water Resources Research* 44, W10403, doi:10.1029/2008WR006848.

936 Green, I.R.A., and D. Stephenson. 1986. Criteria for comparison of single event
937 models. *Hydrological Sciences Journal* 31(3): 395-411.

938 Greig, S.C. 1968. *The geology of the Rigaud Mountain, Quebec.* MSc McGill
939 University, Montreal, Canada. 124 p.

940 Güntner, A., S. Uhlenbrook, J. Seibert, and C. Leibundgut. 1999. Multi-criterial
941 validation of TOPMODEL in a mountainous catchment. *Hydrological Processes*
942 13(11): 1603-1620

943 Harmel, R., R. Cooper, R. Slade, R. Haney, and J. Arnold. 2006. Cumulative uncertainty in
944 measured streamflow and water quality data for small watersheds. *Transactions-*
945 *American Society of Agricultural Engineers* 49(3), 689-701.

946 Hayashi, M., and D.O. Rosenberry. 2002. Effects of ground water exchange on the
947 hydrology and ecology of surface water. *Groundwater* 40(3): 309-316.

948 Henriksen, H.J., L. Troldborg, P. Nyegaard, T.O. Sonnenborg, J.C. Refsgaard, and B.

949 Madsen. 2003. Methodology for construction, calibration and validation of a
950 national hydrological model for Denmark. *Journal of Hydrology* 280(1): 52-71.
951 Hobson, G.D., and J.J.L Tremblay. 1962. *Région cartographiée de Vaudreuil (Québec):*
952 *Partie I - Hydrogéologie de la moitié est et Partie II - Application de la méthode*
953 *séismique pour déterminer la profondeur de la roche en place*. Études 61-20 et Carte 30-
954 1961. Commission géologique du Canada. Ottawa, Canada. 18 p.
955 Hofmann, H.J. 1972. *Excursion B-O3 : Stratigraphie de la région de Montréal*,
956 Montréal, Canada. 34 p.
957 HYDAT. 2014. *Hydrometric Data (HYDAT), Station Sainte-Anne-de-Bellevue*
958 *(02OA013)*. Environment Canada. Retrieved from
959 http://www.wsc.ec.gc.ca/applications/H2O/index-eng.cfm.
960 Hydro-Québec. 2014. *Données 2013-2014 des niveaux d'eau de la rivière des*
961 *Outaouais en aval de la centrale de Carillon*. Unpublished data provided personally.
962 Irvine, D.J., P. Brunner, H.J. Franssen, and C.T. Simmons. 2012. Heterogeneous or
963 homogeneous? Implications of simplifying heterogeneous streambeds in models of
964 losing streams. *Journal of Hydrology* 424–425: 16- 23.
965 Johnson, A.I. 1967. *Specific yield: compilation of specific yields for various materials*.
966 Washington, USA: US Government Printing Office. 74 p.
967 Jones, J.P., E.A. Sudicky, and R.G. McLaren. 2008. Application of a fully- integrated
968 surface-subsurface flow model at the watershed-scale: A case study. *Water Resources*
969 *Research* 44, W03407, doi:10.1029/2006WR005603.
970 Kalbus, E., F. Reinstorf, and M. Schirmer. 2006. Measuring methods for groundwater-
971 surface water interactions: a review. *Hydrology and Earth System Sciences* 10(6):

873-887.

Konikow, L.F., and E. Kendy. 2005. Groundwater depletion: A global problem. *Hydrogeology Journal* 13(1): 317-320.

Kurylyk, B.L., K.T.B. MacQuarrie, T. Linnansaari, R.A. Cunjak, and R.A. Curry. 2015. Preserving, augmenting, and creating cold-water thermal refugia in rivers: Concepts derived from research on the Miramichi River. New Brunswick (Canada). *Ecohydrology* 8(6):1095-1108. doi:10.1002/ eco.1566.

Larocque, M., G. Meyzonnat, M.A. Ouellet, M.H. Graveline, S. Gagné, D. Barnetche, and S. Dorner. 2015. *Projet de connaissance des eaux souterraines de la zone de Vaudreuil-Soulanges : Rapport scientifique.* Report in French submitted to the ministère du Développement durable, de l'Environnement et de la Lutte contre les Changements Climatiques. Montréal, Canada: Université du Québec à Montréal. 189 p.

Lutz, A., J.M. Thomas, G. Pohll, and W.A. McKay. 2007. Groundwater resource sustainability in the Nabogo Basin of Ghana. *Journal of African Earth Sciences* 49(3): 61-70.

McCallum, A.M., M.S. Andersen, G.M. Giambastiani, B.F.J. Kelly, and R.I. Acworth. 2013. River–aquifer interactions in a semi-arid environment stressed by groundwater abstraction. *Hydrological Processes* 27(7): 1072- 1085.

McCuen, R.H. 2004. *Hydrologic analysis and design*. Third Edition. Upper Saddle River, New Jersey, USA: Prentice Hall. 888 p.

MDDELCC (Ministère du Développement durable, de l'Environnement et de la Lutte contre les changements climatiques). 2013. *Système d'information hydrogéologique (SIH) – Well drillers' database.* http://www.mddelcc.gouv.qc.ca/eau/souterraines/sih/ accessed in November 2013.

MDDELCC (Ministère du Développement durable, de l'Environnement et de la Lutte contre les changements climatiques). 2014. *Climatologie. Direction du suivi de l'état de l'environnement.* Québec, Canada: Ministère du Développement durable, de l'Environnement et de la Lutte contre les changements climatiques. http://www.mddelcc.gouv.qc.ca/climat/surveillance/index.asp

Moreira, F. 2016. *Estimation de la décharge des eaux souterraines dans deux rivières du Québec par le traçage du ^{222}Rn et de l'argon*. MSc thesis in French, Université du Québec à Montréal. 122 p.

Nash, J., and J. Sutcliffe. 1970. River flow forecasting through conceptual models part I-A discussion of principles. *Journal of Hydrology* 10(3): 282-290.

Pelletier, P.M. 1988. Uncertainties in the single determination of river discharge: a literature review. *Canadian Journal of Civil Engineering* 15(5), 834-850.

Québec-Pêche. 2011. *Forum: cartes Navionics.* http://www.quebecpeche.com/forums/index.php?/topic/107701-cartes-navionics/

Refsgaard, J.C. 1997. Parameterisation, calibration and validation of distributed hydrological models. *Journal of Hydrology* 198(1-4): 69-97.

Roy, M., and P.M. Godbout. 2014. *Cartographie des formations superficielles de la région de Vaudreuil-Soulanges, sud-ouest du Québec*. Report in French submitted

1015 to the Ministère des Ressources Naturelles du Québec, Université du Québec à
1016 Montréal, Montréal, Canada. 25 p.
1017 Rozemeijer, J.C., and H.P. Broers. 2007. The groundwater contribution to surface
1018 water contamination in a region with intensive agricultural land use (Noord-
1019 Brabant, The Netherlands). *Environmental Pollution* 148(3): 695-706.
1020 Rutqvist, J., J. Noorishad, C.F. Tsang, and O. Stephansson. 1998. Determination of
1021 fracture storativity in hard rocks using high-pressure injection testing. *Water*
1022 *Resources Research* 34(10), 2551-2560
1023 Sahoo, G.B., C. Ray, and E.H. De Carlo. 2006. Calibration and validation of a
1024 physically distributed hydrological model, MIKE SHE, to predict streamflow at
1025 high frequency in a flashy mountainous Hawaii stream. *Journal of Hydrology*
1026 327(1–2): 94-109.
1027 Sauer, V.B. and R. Meyer. 1992. *Determination of error in individual discharge*
1028 *measurements*. Open-file report no. 92-144. Norcross, Georgia, USA: US Department of
1029 the Interior, US Geological Survey. 21 p.
1030 Spanoudaki, K., A.I. Stamou, and A. Nanou-Giannarou. 2009. Development and
1031 verification of a 3-D integrated surface water–groundwater model. *Journal of*
1032 *Hydrology* 375(3–4): 410-427.
1033 Sulis, M., Paniconi, C., et Camporese, M. (2011). Impact of grid resolution on the
1034 integrated and distributed response of a coupled surface-subsurface hydrological
1035 model for the des Anglais catchment, Quebec. *Hydrological Processes, 25*(12),
1036 1853-1865

1037 Technorem. 2005. *Élaboration d'un mode de gestion et d'exploitation du système*
1038 *aquifère de la ville de Hudson*. Report no. PR04-50. Laval, Canada: Technorem
1039 Inc. 463 p.
1040 Teloglou, I.S., and R.K. Bansal. 2012. Transient solution for stream–unconfined
1041 aquifer interaction due to time varying stream head and in the presence of
1042 leakage. *Journal of Hydrology* 428–429: 68-79.
1043 Todd, D.K. 1980. Groundwater hydrology. 2nd Édition. New York, USA: Wiley.
1044 535 p.

1045 Vazquez, R.F., L. Feyen, J. Feyen, and J.C. Refsgaard. 2002. Effect of grid size on
1046 effective parameters and model performance of the MIKE-SHE code.
1047 *Hydrological Processes* 16(2): 355-372.
1048 Williams, J.H., R.J. Reynolds, D.A. Franzi, E.A. Romanowicz, and F.L. Paillet. 2010.
1049 Hydrogeology of the Potsdam Sandstone in northern New York. *Canadian Water*
1050 *Resources Journal* 35(4): 399-416.
1051 Winter, T.C., J.W. Harvey, O. Franke, and W. Alley. 1998. *Ground water and*
1052 *surface water: a single resource*. US Geological Survey circular 1139. Denver,
1053 Colorado, USA: USGS. 87 p.
1054 Zhou, X., M. Helmers, and Z. Qi. 2013. Modeling of Subsurface Tile Drainage using
1055 MIKE SHE. *Applied Engineering in Agriculture* 29(6): 865-873

1056

1057

1058 **Table 1. Measured and calibrated parameters**

	Calibrated	Range of measured values
Hydraulic conductivity[(1)] (*K*) (m.s^{-1})		
Sand	3.0×10^{-5}	7.0×10^{-6} - 7.4×10^{-3}
Till	1.0×10^{-6}	3.5×10^{-7} - 1.1×10^{-4}
Clay	5.0×10^{-8}	1.6×10^{-10} - 9.7×10^{-8}
Crystalline bedrock	1.6×10^{-7}	1.0×10^{-6} - 8.3×10^{-5}
Sedimentary bedrock	1.4×10^{-4}	1.1×10^{-7} – 9.0×10^{-4}
Storage coefficient[(2)] (*S*) (-)		
Sand	0.25	0.10 - 0.35
Till	0.15	0.06 - 0.16
Clay	0.025	0.00 - 0.05
Crystalline bedrock	0.01	0.01 - 0.05
Sedimentary bedrock	0.05	0.18 - 0.30
Specific storage[(3)] (S_s) (m^{-1})		
Sand	3.0×10^{-4}	1.0×10^{-5} - 1.0×10^{-4}
Till	1.0×10^{-4}	1.0×10^{-5} - 1.0×10^{-3}
Clay	1.0×10^{-4}	1.0×10^{-4} - 1.0×10^{-2}
Crystalline bedrock	1.0×10^{-5}	1.0×10^{-7} - 1.0×10^{-5}
Sedimentary bedrock	1.0×10^{-6}	1.0×10^{-7} - 1.0×10^{-5}
River leakage coefficient[(4)] (r_l) (s^{-1})		
Raquette River and main tributaries	7×10^{-7}	3.9×10^{-7} - 1.4×10^{-3}
Drainage constant[(5)] (C_d) (s^{-1})		
Small streams	5×10^{-6}	2.0×10^{-8} - 4.9×10^{-4}
Manning coefficient[(6)] (*n*) (m$^{1/3}$.s^{-1})		
Agricultural area (clay)	15	1.7 - 83.3
Infiltration fraction[(7)] (f_I)		
Sand	0.70	0.44 - 0.66
Till	0.42	0.33 - 0.51
Clay	0.05	0.03 - 0.07
Undifferentiated bedrock	0.60	0.66 - 0.89

1059 [(1)] Measured values reported in Larocque et al. (2015).
1060 [(2)] Johnson (1967); Todd (1980)
1061 [(3)] Rutqvist et al. (1998); Domenico and Mifflin (1965)
1062 [(4)] Doppler et al. (2007); Spanoudaki et al. (2009); Irvine et al. (2012); Teloglou and
1063 Bansal (2012)
1064 [(5)] Al-Khudhairy et al. (1999); DHI (2007); Zhou et al. (2013)
1065 [(6)] McCuen (2004)

1066 [7] Croteau et al. (2010)

Figure captions

Figure 1. Location of the Raquette River watershed in the Vaudreuil-Soulanges region of southern Quebec (Canada), including topography and monitoring stations.

Figure 2. Geological and hydrogeological conditions in the Raquette River watershed; a) Quaternary deposits (modified from Roy and Godbout 2014), and b) piezometric map (modified from Larocque et al. 2015).

Figure 3. Measured flow rates at 23 locations along the Raquette River on August12th and 25th 2015. The black arrows indicate tributaries 1 (T1), 2 (T2), and 3 (T3). The grey arrows indicate the main streams. The dashed vertical lines indicate the locations of gauging stations 1, 2, and 3.

Figure 4. Net monthly vertical inflows (VI_{net}) and air temperature (Temp.) in the Raquette River watershed for the two hydrological years studied (Nov. 2012 – Oct. 2013 and Nov. 2013 – Oct. 2014). Vertical inflows (VI) correspond to the sum of precipitation occurring as rain when air temperature exceeds the freezing point and snowmelt occurring during the winter and in the spring snowmelt period. VI_{net} corresponds to VI values from which evapotranspiration has been substracted. VI_{net_cal} corresponds to the calibrated VI_{net} values.

Figure 5. Boundary conditions used to represent the Raquette River study area in the integrated MikeSHE model.

Figure 6. Absolute values of total sensitivity coefficients (S_{TOT}) with respect to a) flow rates, and b) heads, corresponding to 5% parameter variation. Plus signs indicate a positive sensitivity, whereas, whereas no sign indicates a negative sensitivity.

Figure 7. Measured and simulated flow rates at gauging stations a) 1, b) 2, and c) 3.

Figure 8. Simulated flow rates at the outlets of tributaries 1, 2, and 3.

Figure 9. Measured and simulated steady-state heads. The thick black line is the 1:1 line and the dotted lines represent a +/- 5 m error envelope.

Figure 10. Measured and simulated transient-state heads at the five observation wells. Note that the vertical scale of the PO4 crystalline bedrock well is larger than that of the other wells.

Figure 11. Monthly simulated net vertical inflows (VI_{net}) and flows between gauging stations 1 and 2, between gauging stations 2 and 3, and above gauging station 3 for a) drained flows and b) baseflows.

REVIEWER 1'S COMMENTS:

GENERAL COMMENTS

This is a case study in a 133 km^2 watershed in Quebec, where a commercial model MIKE-SHE is used to simulate river discharge and aquifer water levels. The authors use relatively simple approaches to estimate net water inputs to the system and to characterize aquifer properties. The objective of the study is to identify the challenges encountered during model setup and calibration using a limited data set. The manuscript is well written in a clear language and technical contents are sound. It is a good documentation of how to set up the MIKE-SHE model in data-limited environment and evaluate its performance. However, as it is written, the manuscript does not clearly demonstrate the significance and relevance of the study in the context of water resources evaluation and management in Canada and elsewhere. Models like MIKE-SHE are routinely used by consultancy and regulatory agencies in glaciated and non-glaciated environments. It is desirable to see the relevance of the case study in a broader hydrogeological context. For example, what is the intended use of the model? Given that, are the uncertainties and discrepancies in the model acceptable? How are the model parameters determined? How sensitive is the model to the variability in some of the critical parameters? Addressing these questions will strengthen the paper and make it more interesting and useful to the readers of this journal

We thank the reviewer for his thorough and constructive comments. We have attempted to address them all as best as possible, and we think that the new version of the manuscript is now much improved.

We agree that integrated models are increasingly used by consultants and regulatory agencies around the world. However, to the best of our knowledge, these modelling applications are not yet routinely used, due either to the lack of available data, or to the lack of human and computer resources needed to develop the models. A thorough analysis of integrated flow conditions, including a sensitivity analysis and an examination of the areas where improvement is necessary to build a more robust model, is usually reserved to academic contexts. This paper addresses both of these points (and we thank the reviewers for suggesting that a sensitivity analysis be carried out) in an attempt to make the exercise useful for non-academic users. Integrated models that are used and improved over the long-term to guide integrated water management are still rare (these sentences were added L89-98). This paper is intended to provide guidelines to identify the challenges related to this goal and to contribute to the development of models that will become truly useful water management tools.

This modelling exercise has allowed many uncertainties in the model results to be identified. Although these uncertainties are relatively high (especially for heads), the integrated model still provides useful information. As mentioned in the text, the model parameters were manually calibrated through a trial-and-error process (see the response to comment 19). Combined with the sentisitivity analysis, identification of the processes that are less well represented by the model helped to identify what the critical parameters are and how model improvement could be achieved. This is now discussed in the new version of the manuscript (L515-534 and L634-664).

SPECIFIC COMMENTS

1. *Line 125. Grid, instead of matrix, is more commonly used in the literature.*

We agree with the reviewer that "grid" is more adequate and have made the change as suggested (L130).

2. *Figure 2. Please indicate the surface drainage divide in the map. The legend in Fig. 2b says "water table elevation", but the caption says it is "piezometric map". What is actually shown? Piezometeric head or the water table?*

As suggested, we have added the surface drainage divide to the two pannels of Figure 2. We have also replaced "watertable elevation countours" with "piezometric head" in the legend of Figure 2b.

3. *Line 213. No hydraulic connection. This is not clear in the map. Please explain.*

The piezometric map shows that the Raquette River drains the aquifer in its downstream portion, as indicated by the arrows (some arrows were added for clarification). In the upstream portion of the Raquette River (corresponding to the clay plain), the piezometric map is not influenced by the river channel, indicating that there is no hydraulic connection between the confined aquifer and the river. We have modified the text to clarify this (L273-275).

4. *Line 219-220. Heads were constrained to river water levels. What does this exactly mean? Please provide a more specific explanation.*

We agree with the reviewer that the forcing of the piezometric map where the river is known to drain the aquifer was not clear. We have revised the text to explain that between the Saint-Lazare and Hudson hills, head control points equal to river water levels were used in the interpolation to compensate for the small number of measured heads (see Figure 1) and to reproduce the drainage effect of the river in the piezometric map (L280-283).

5. *Line 229. What is the averaging period of annual precipitation (e.g. 1981-2010)? Were winter precipitation values adjusted for wind-induced undercatch?*

We thank the reviewer for identifying that the reference period (1981-2010) was missing. This information was added to the new version of the manuscript (L289). Winter snow precipitation values were not adjusted for wind-induced undercatch or sublimation. Although this can be an important factor in some conditions, very few data on this process are available in the literature. We have added a sentence to the *Methods* section, mentioning the existence of wind-induced undercatch and sublimation as processes that may remove snow from the local water budget, but were not included in this work (L299-302). However, because the simulated flow rates are lower than the measured values, it is hypothesized that these processes do not reduce significantly the water available at snowmelt. For this reason, we did not consider relevant to discuss the possible impact of not taking them into consideration in the quality of the model results.

6. *Line 231. What does "respectively" refer to? 2013 and 2014? If the hydrological year is different from the calendar year, please define it here.*

We agree with the reviewer that the sentence was poorly constructed. The text was revised to clarify that a total precipitation of 1049 mm/yr was observed between November 2012 and October 2013, and a total precipitation of 1119 mm/yr was observed between November 2013 and October 2014 (L291-293).

7. *Figure 3. Please explain the acronyms (VI_net_cal etc.) in the figure caption. The caption says "VI is precipitation plus snowmelt". Does this mean that snow is double counted, first as snowfall and second as snowmelt? Please clarify in the texts.*

As suggested by the reviewer, we have added explanations of the acronyms VI_{net}, VI_{net_cal}, and Temp. to the caption of Figure 4 (L1054-1059; previously Figure 3). We have also added a description of *VI* (i.e., the sum of precipitation occurring as rain when air temperature is above the freezing point and snowmelt occurring during the winter and in the spring snowmelt period), and have also modified the main text to clarify this definition (L307-309).

8. *Line 247. A substantial portion of snow can be lost to interception and sublimation. How are these processes accounted for in the model?*

As mentioned in response to comment 5, wind-induced undercatch and snow sublimation were not taken into consideration in the model.

9. Line 248. What does "global" mean? Does this mean a global (whole earth) scale hydrological model? Please clarify.

We have changed the sentence to remove the term "global" (L312-314).

***10.** Line 269. How frequently was the stream discharge measured? What was the uncertainty (e.g. RMSE) of the rating curve? This information is critical, as it related to the interpretation of model results later.*

We have followed the reviewer's suggestion and included the number of flow rate measurements taken at the three gauging stations (23, 18, and 26 for stations 1, 2 and 3 respectively; L205-209). We have also added the RMSE for each of the three stations (L206-208).

***11.** Line 279. Were there no other hydrological features capable of water retention, such as wetlands and lakes, in the watershed?*

There are very few wetlands and no lakes capable of water retention in the watershed. We agree with the reviewer that it is important to mention this, and we have added a sentence at the end of the *Physiography and hydrography* section (L146-147).

***12.** Line 287. Please indicate the extent of the model domain in Figure 4.*

There was an error in the figure (previously Figure 4, now Figure 5) of the submitted version. The corrected figure has now been included, and describes the model boundary conditions.

***13.** Figure 4. Please indicate the surface drainage divide. Contour lines and intermittent linear topography are indistinguishable in the figure. Please use different line type or thickness.*

See our response to comment 12 above.

***14.** Line 312. Should this be horizontal, not vertical? How was 10% chosen?*

We thank the reviewer for noticing this error: K_v is 10% of K_h (modified L352-354). A ratio of 10% is freqently used when no measured values are available.

***15.** Line 329. Why was 25.4 mm (one inch) chosen for this watershed? Was this parameter included in model calibration?*

Detention storage was set to the default value in MikeSHE, of 25.4 mm (one inch). This process is very difficult to measure at the watershed scale. It was not tested in the sensitivity analysis, because no reasonable boundary values are available.

***16.** Line 339. It is not clear how surface roughness coefficient can represent subsurface tile drainage in the model. Please provide a convincing explanation.*

We agree with the reviewer that this was not well explained. This watershed (as is the case for most watersheds in agricultural areas of southern Québec) has intensive tile drainage. These artificial drains intercept infiltrating water and redirect it toward the Raquette River, thus accelerating floods and shortening hydrogram response times. The Manning roughness coefficient was calibrated, because it can provide a similar effect on river hydrograms to that of tile drainage. We have rephrased the text explaining this (L377-381), and we have discussed this further in the *Discussion* section (L493-498).

***17.** Line 354. I do not understand "boundary imposed by groundwater flow directions". Please explain.*

We agree with the reviewer that this sentence was unclear. What was meant was that the other model boundaries were flow boundaries, coinciding with either the watershed boundary (no-flow) or with boundaries imposed by groundwater flow directions (imposed flux). The text has been modified to clarify this (L394-396).

18. *Line 367. I do not understand "repeated five times in the simulation to ensure convergence". Please rephrase.*

We have clarified the text to address this ambiguity (L407-409). A ten-year spin-up period, in which the meteorological data for the two years of the study period were repeated five times, was used to ensure that initial conditions were stable before simulating the actual conditions.

19. *Line 370. Calibration by trial-and-error. This is where the study can be significantly improved. It is becoming the standard practice to use a systematic calibration approach when a model has a relatively small number of parameters (only seven in this case), for example PEST (http://pesthomepage.org/), GLUE (generalized likelihood uncertainty estimation), or other algorithms. Since this study is focused on model development and calibration, it is highly desirable to use a more rigorous calibration approach.*

We agree with the reviewer that performing an automatic calibration would have been a reasonable approach to start with. Using an automatic calibration tool such as PEST would have provided further insight into the modelling process. However, the current model uses 22 parameters (not seven as mentioned by the reviewer), and this large number of parameters can lead to parameter non-unicity even with an automatic procedure. We have therefore decided to retain the manual calibration of the model. We believe that this approach does not limit the insight provided by the results. However, we have highlighted that a different parameter combination might have been obtained with an automatic calibration procedure (L661-664).

20. *Line 374. For a model testing study like this, it is important to include the model sensitivity analysis. How sensitive are the simulated stream flows and hydraulic heads to the variability in each of the seven parameters? Is the model particularly sensitive to some of the parameters? If so, what are the consequences?*

This is a very good point. We have followed the reviewer's recommendation and have performed a sensitivity analysis of the model to its calibrated parameters. A description of the method used to calculate the model sensitivity is now included in the *Methods* section (L457-471). The results are presented with a new figure (Figure 6) and they are described in the *Results-Calibrated parameters* section (L515-534). This has allowed us to better understand to which parameters the heads and the flow rates are the most sensitive, and to underline where parameterization efforts should be invested in this type of integrated model. A separate sub-section is now dedicated entirely to model sensitivity at the beginning of the *Discussion* section (L633-664).

21. *Lines 414-416. Please delete periods between m3 and s-1. This applies to the rest of the paper.*

We made the requested modification throughout the paper.

22. *Line 418. "Relatively" is redundant. Please delete.*

"Relatively" was removed as suggested.

23. *Line 471. Why half? Please explain.*

The percentage of reduction in winter available precipitation that was transferred to available water in the spring resulted from a manual trial-and-error process. This explanation was added to the new version of the manuscript (L506-514).

24. *Line 473. Please see my comment on Line 247 regarding snow loss.*

As mentioned above, wind-induced undercatch and snow sublimitation were not taken into consideration in the model.

25. *Line 514. If I understood correctly, the 149 heads were measured only once. What is the expected temporal variability of hydraulic heads in these wells? How does that affect the interpretation of Figure 8?*

The reviewer is correct, 149 heads were measured in private wells, only once. Heads were also available from five monitored observation wells. The text at L258-259 and L260-261 was modified to clarify this. Measurements from the monitored observations show that the heads had

amplitudes (maximum minus minimum measured head) ranging from 0.29 m (CPT1) to 6.39 m (PO4). From these values, it is clear that head variations are greatest at the highest topography in the crystalline bedrock of Mount Rigaud (L261-265). Error bars were not included in Figure 8 (now renamed Figure 9), because they would only have been visible on the graph for well PO4. However, we have included a more complete description of errors on simulated heads, by comparing the MAE to the maximum amplitude measured at well PO4 (L574-578).

26. *Line 517. I would say that RMSE of 8.77 m is quite large. Please provide a justification for considering this an "acceptable" value.*

This interpretation is based on our experience in groundwater simulation. An RMSE of 8.77 m appears to be reasonable in conditions where there is a large topographic gradient. The normalized RMSE, which puts this into perspective, is lower than the 0.1 target usually considered to be acceptable (L578-579). However, we acknowledge in the text that the largest simulated errors are not located at the highest topography, on Mount Rigaud, but rather at the lower elevation Hudson Hill, and at the eastern and northern extents of Mount Rigaud (L579-583). Possible causes for these errors are discussed in L724-738.

27. *Line 519-520. It will be useful to include a figure indicating the spatial variability of errors.*

We do not think that this figure would add useful information to the paper. We believe that the piezometric map (Figure 2b) and the scattergram of measured and simulated heads (formerly Figure 8, now Figure 9) provide sufficient information for the reader to locate where the largest errors are in the study area.

28. *Figure 10. Please explain in the caption that these are simulated values.*

We have made the suggested modification (L1072) in the caption of Figure 11 (formerly Figure 10).

29. *Line 580. The uncertainty in rating curve. This is a very important topic and warrants more detailed discussion. If possible, please present the rating curves and discuss how the overestimation of high flow (which occurs only for a short period) can result in a large overestimation of total annual flow.*

We agree with the reviewer that the uncertainty in the rating curve is an important topic in the discussion. As suggested above by the reviewer, we have included the RMSE on the measured flow rates to estimate the uncertainty in the rating curves (L206-208). This uncertainty is linked to the estimation of high flows, and is especially significant at station 1 (L209-210). The *Discussion-Simulated flow rates and heads* discusses that the overestimation the apparent overestimation of measured total flows compared with simulated values could be due to the imprecision of the rating curves (L673-674). However, it is also mentionned that this overestimation could be due to the simplified representation of evapotranspiration and infiltration processes (L670-673). We have chosen not to include the rating curves themselves to avoid making the paper any longer.

30. *Line 734. What are the specific "interests for water management"? How will the model be used by water managers? Will it be still useful, given the uncertainties and limitations?*

We agree with the reviewer that this sentence was too vague. We have reformulated the end of the conclusion (L206-208) to explain how water managers are increasingly asked to implement integrated water management and to develop adaptation measures to changing land use and climate conditions. We have also included one sentence in the introduction to explain how water managers can use this type of integrated model (L98-100). A new *Recommandation* section was added to summarize the results from this study (L798-828). These recommandations are intended for the development of integrated water flow models by water scientists, but also for the use of such models by water managers.

REVIEWER 2'S COMMENTS:

GENERAL COMMENTS

This manuscript presents an application of MIKE-SHE hydrological model to the Raquette River watershed in Quebec, Canada. The model was calibrated and compared against flow rate and groundwater head measurements at different points of the study area. Authors performed simulations covering two hydrological years with comparison and discussion limited to snow-free conditions.

I think that accurate modeling efforts supported by observations are always important exercises that have the potential of providing useful information to the scientific community. In so doing such model applications should identify some clear research issues that are addressed throughout the manuscript. This work shows, in my opinion, some of these potentials (e.g., sound model setup, valuable hydrological and hydrogeological characterization, good set of observations) with some clear limitations in identifying a tangible outcome of their modeling exercise for other users. This issue requires, in my opinion, major re-structuring of the manuscript with the inclusion of some additional results. I provide below some more specific comments that support my general idea.

We agree with the reviewer that the previous version of the manuscript did not clearly identify a tangible outcome of the modeling exercise. We have addressed this in the introduction of the current version by explaining that, although integrated models are increasingly used by consultants and regulatory agencies, they are not yet routinely applied for many reasons (i.e., lack of data; lack of human and technical resources). Furthermore, when integrated models are used, their application is not necessarily complete (L92-98). This paper aims to identify the common challenges encountered in glaciated terrains with regards to parameterizing and simulating integrated flow conditions. Through a sensitivity analysis, and with a discussion of model uncertainties, this work pinpoints possible model improvements that should be targeted to develop a long-term integrated water management tool (see additions to the *Discussion* section and the new *Recommandations* section L798-828). We thank the reviewer for suggesting to perform a sensitivity analysis of the model. This exercise has allowed us to include what we believe are significant new results that broaden the scope of the paper's contribution.

SPECIFIC COMMENTS

1. *Authors set an ambitious objective for their work, that is, "assess" when fully-coupled models are useful and "how" they should parameterized (see lines 24-25). I think that results presented in the manuscript do not shed lights on these critical points. To proof these statements you need a more robust framework than the one you have in the manuscript. My concern is somehow testified by the sentence at lines 87-90 where authors try to identify the need of their work. Which method are you talking about? I think this a delicate issue where authors should make a substantial effort in re-thinking and re-structuring their effort.*

The *Introduction* illustrates the broad context underlying the necessity of examining the application of integrated surface and groundwater flow models. We agree with the reviewer that the previous version of the manuscript may have suggested that the paper would address some issues that were , in fact, beyond its scope. We have modified the *Introduction* to more realistically present the context of the work reported here. The beginning of the *Abstract* was revised to better contextualize that the paper aims to guide the implementation of integrated models (L23-25). The objective of the paper was rephrased as follows: "to examine some of the challenges encountered when simulating surface water – groundwater interactions in a post-glacial geological environment" (L115-116). We believe that this slightly modified version of the objective is more coherent with the actual scope of the paper.

2. In a similar vein to the previous point, authors make a series of simplifying hypothesis in setting up the model to arrive at the conclusion that these simplifications were potentially at the origin of the mismatch between simulated and measured results. It is not clear to me what we can really learn (and use for future studies) from this modeling exercise because we're still in the realm of hypothesis. A better approach would have implied the testing (with different model configurations) of at least one such simplification. This is the case where good modeling applications convey a strong take-home message. This should be addressed presenting more results in the revised version of the manuscript.

We agree with the reviewer that testing different model configurations would have provided a very different perspective for this paper. Although such an exercise would have been very instructive, it was beyond the scope of the application of this model, and thus could not be implemented in the new version of the manuscript. Nevertheless, we believe that the sensitivity analysis that was added to the manuscript brings some insight as to which parameters (and processes) dominate the flow calculations.

3. The organization of the manuscript could be improved moving the presentation of the measurements outside the "Results" section. Similarly I would include the calibration exercise within the methodological part of the manuscript as this step consists mainly in finding a good setup of the model.

We agree with the reviewer that the description of flow rates is better located before the *Results* section. We have therefore added a new *Study area-Hydrology* section (L184-238) where we have transferred a simplified version of the hydrology measurements and the full flow rate value description that was previously in the *Results* section. This modification required us to change the order of Figures 3, 4, and 5.

However, we believe that the results from the calibration should remain in the *Results* section, especially since we have performed a sensitivity analysis in MikeSHE. This is now an important contribution from this study.

4. Some improvements are needed in the language and terminology used in the manuscript. An example is given at lines 102-108. Here the language should be, in my opinion, elevated a bit more. What do you mean here with "tedious"? What do you mean with the sentence at line 367? What convergence we are talking about? Moreover, authors use throughout the paper the term "fully-coupled" but this is not really the type of modeling approach they implemented for their case study.

We have seriously considered the reviewer's comments concerning the language and terminology. We have reviewed the entire paper to bring the text to the required level of scientific English. In particular, we have rephrased the passage that made use of the word "tedious". We have also reviewed the sentence at line 367 to more clearly explain how the model used a spin-up period (L407-409; also see our repsonse to comment 18 from Reviewer 1). The term "fully-coupled" used throughout the previous version of the manuscript has been replaced by the more accurate "integrated" modelling approach in the new version. Prior to this new submission of the paper, we also had an English speaking scientist review the text entirely to provide the best possible level of language (see version of the manuscript with track mode for changes)

ASSOCIATE EDITOR'S COMMENTS:

I can now inform you that the reviewers (two) have evaluated your manuscript entitled "Examining the challenges of simulating surface water – groundwater interactions in a post-glacial environment". Both reviewers have issues and comments to consider for improving the manuscript, particularly regarding the calibration approach and the need to clearly demonstrate the outcome and the relevance of the case study. Please, along with the comments of Reviewer #1 and Reviewer #2, you should make the following corrections to Figure 9 (Dates should be moved to the bottom x-axis so that they do not overlap the lines) and Figure 10 (Lines should not overlap text in Figure legend).

We have addressed all of the reviewers' and the associate editor's comments to the best of our ability. In particular, we have better defined the aim of the paper, and more clearly demonstrated the relevance of this case study for research, consultants, and regulatory agencies. Although we have chosen to retain the manual calibration approach, the inclusion of a sensitivity analysis now provides additional content that broadens the scope of the paper (new Figure 6 and new text in the *Methods*, *Results*, *Discussion*, and *Recommandations* sections). We have incorporated the modifications requested by the associate editor for Figures 10 and 11 (formerly Figures 9 and 10). We believe that the new version of the manuscript is now substantially improved, and we hope that it satisfactorily responds to any concerns of the reviewers and the associate editor.

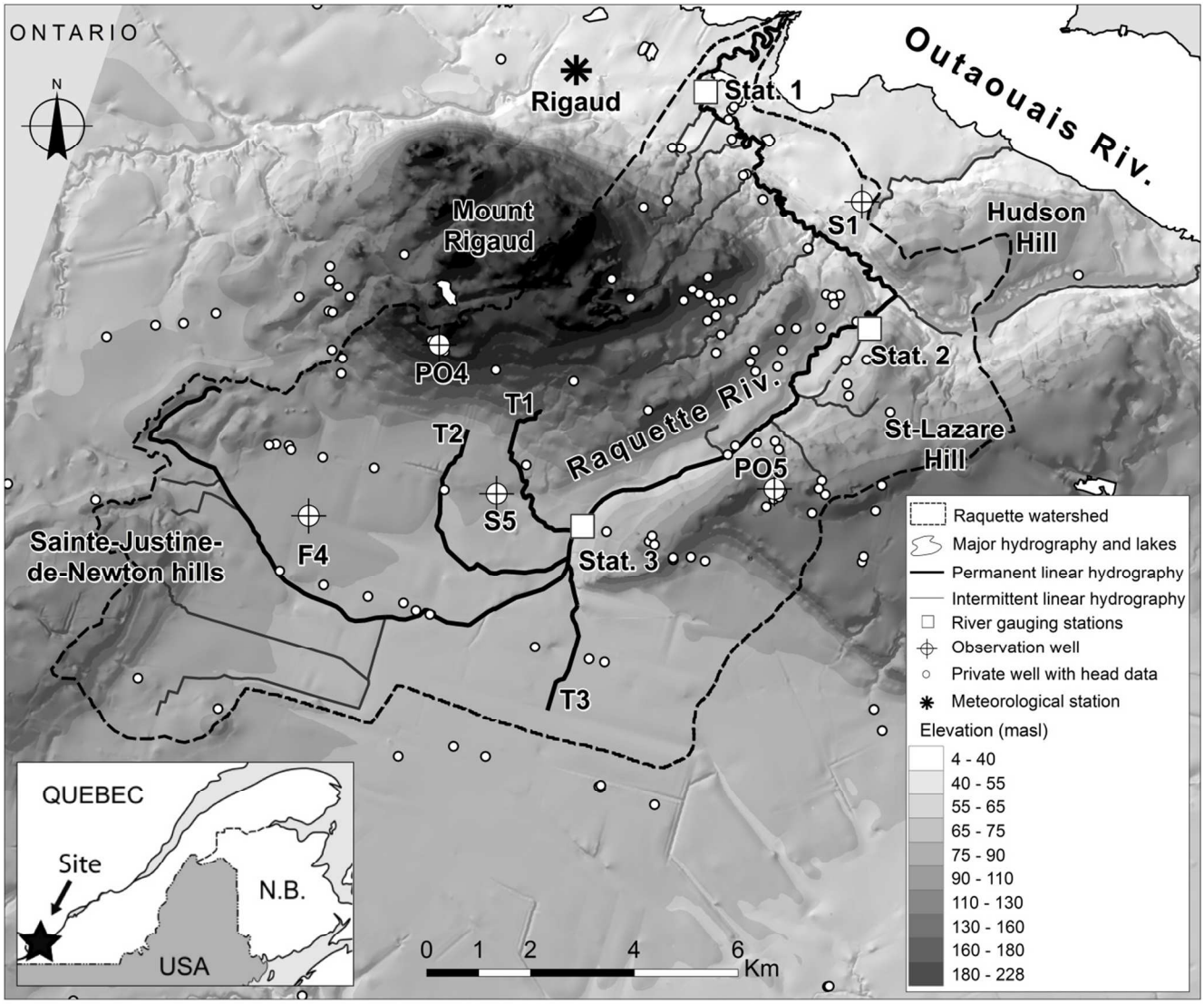

Figure 1.
Turgeon et al 2017
Submitted to Canadian Water Resources Journal

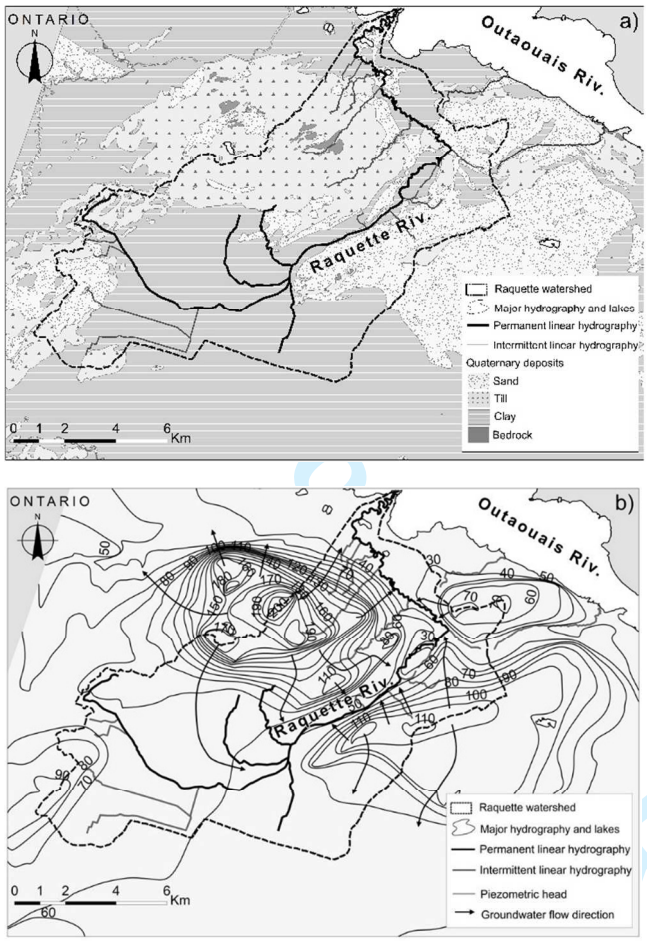

Figure 2.
Turgeon et al 2017
Submitted to Canadian Water Resources Journal

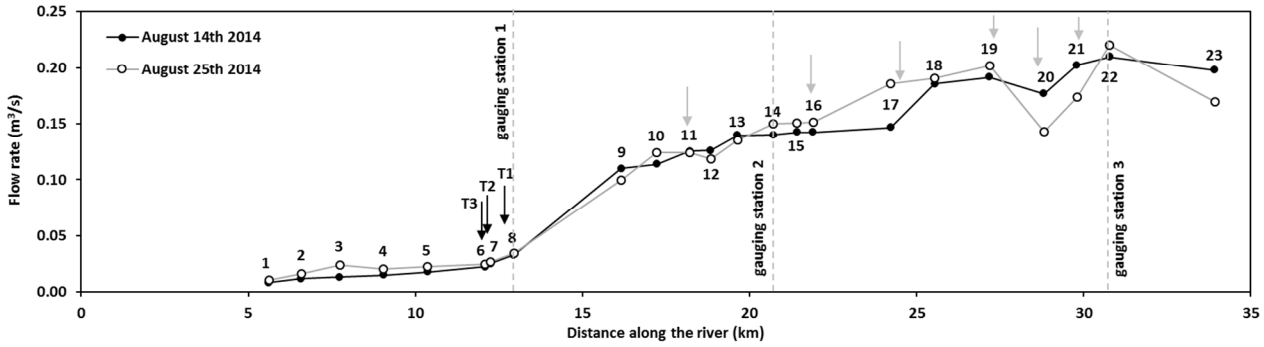

Figure 3.
Turgeon et al 2016
Submitted to Canadian Water Resources Journal

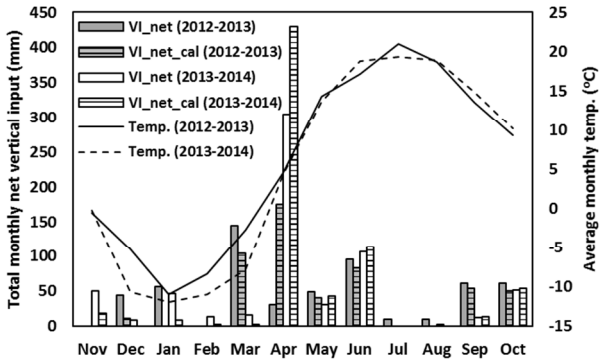

Figure 4.
Turgeon et al 2016
Submitted to Canadian Water Resources Journal

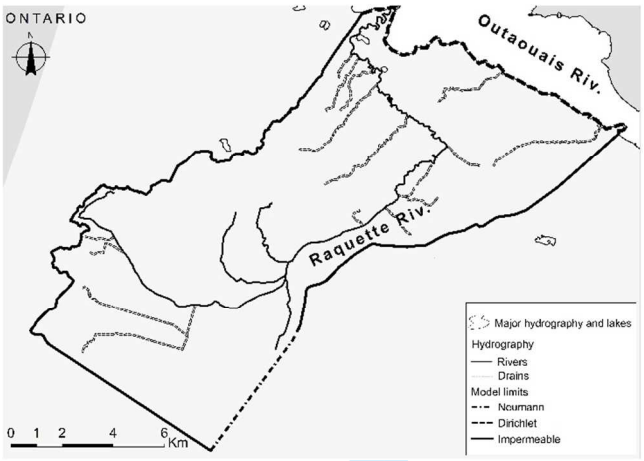

Figure 5.
Turgeon et al 2017
Submitted to Canadian Water Resources Journal

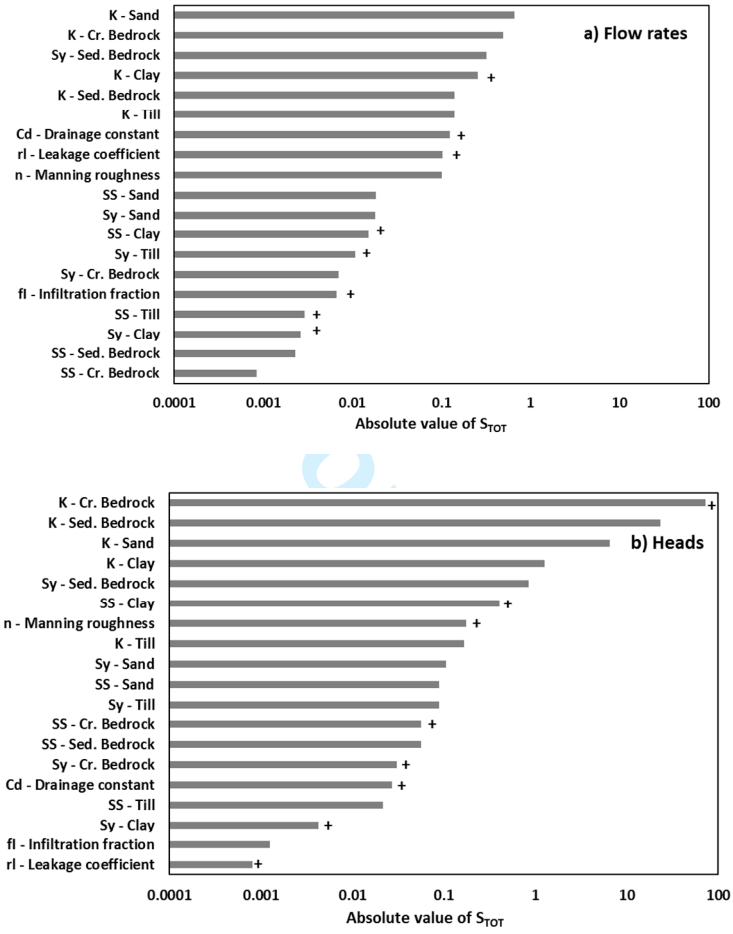

Figure 6.
Turgeon et al 2017
Submitted to Canadian Water Resources Journal

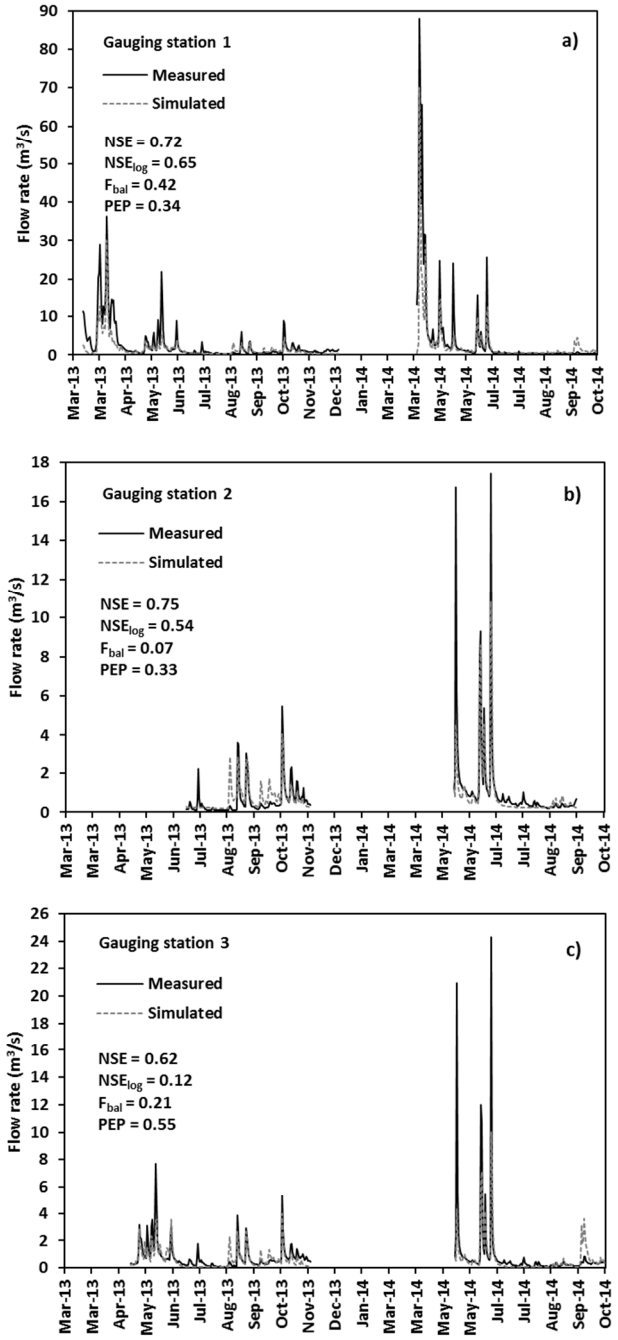

Figure 7.
Turgeon et al 2017
Submitted to Canadian Water Resources Journal

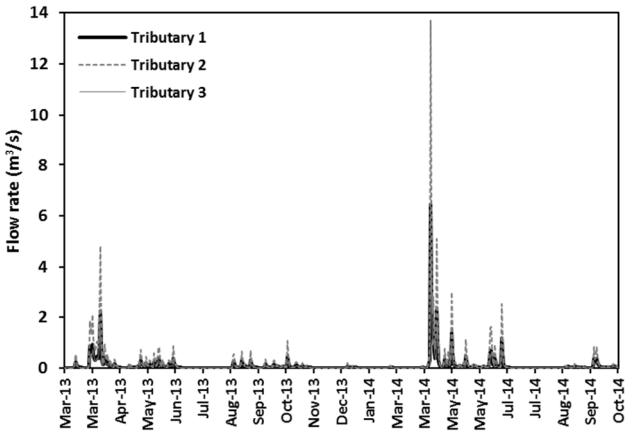

Figure 8.
Turgeon et al 2017
Submitted to Canadian Water Resources Journal

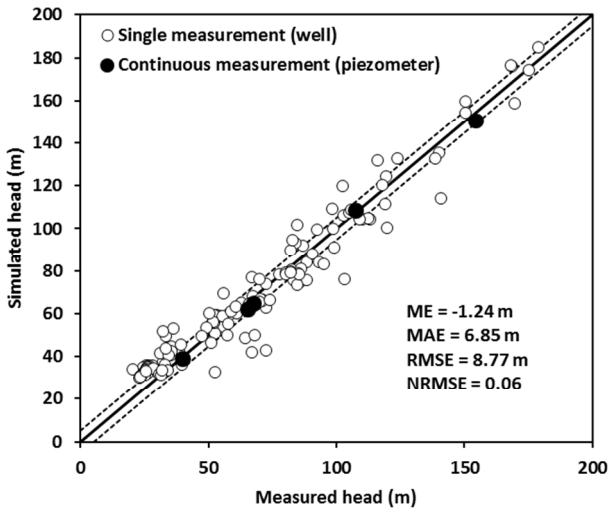

Figure 9.
Turgeon et al 2017
Submitted to Canadian Water Resources Journal

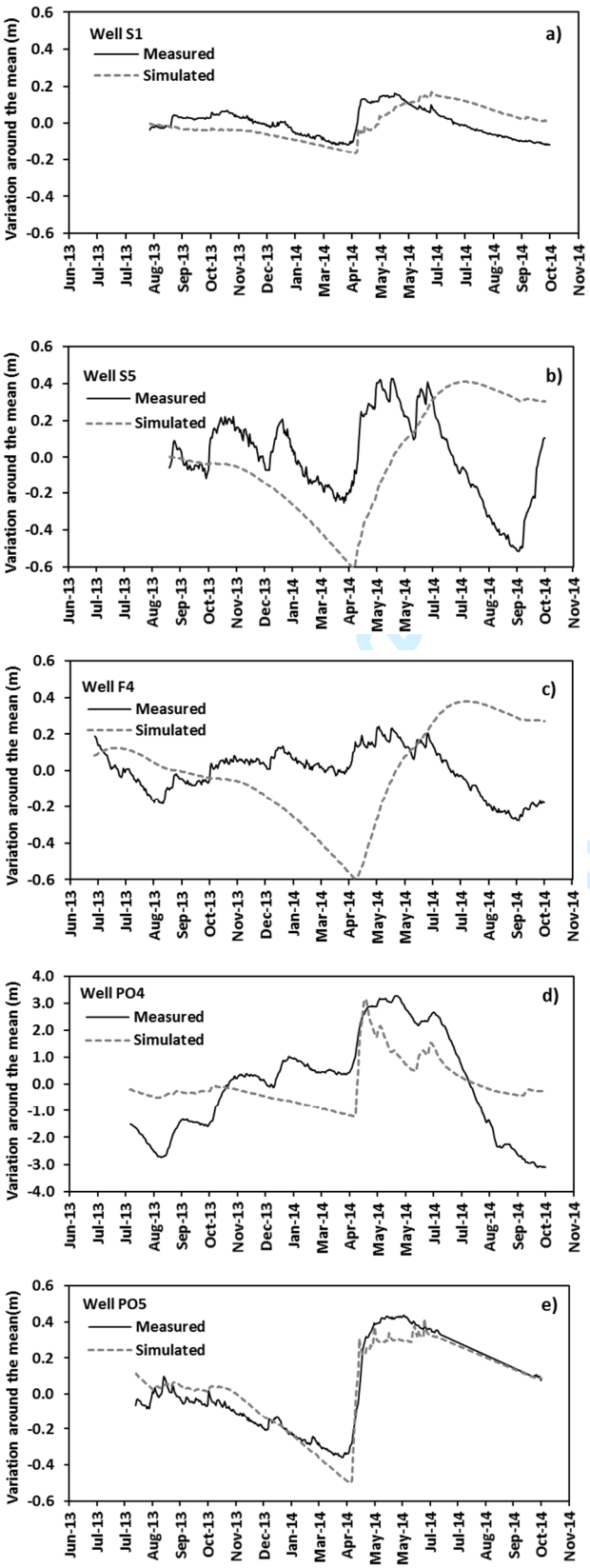

Figure 10.
Turgeon et al 2017
Submitted to Canadian Water Resources Journal

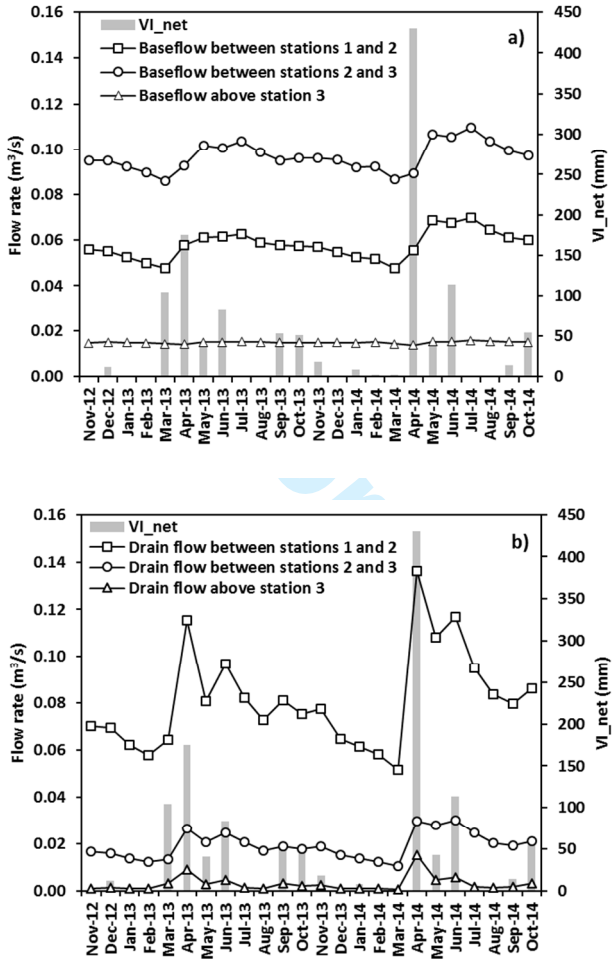

Figure 11.
Turgeon et al 2017
Submitted to Canadian Water Resources Journal