Assimilation of passive microwave AMSR-2 satellite observations in a snowpack evolution model over northeastern Canada

Fanny Larue1,2,3, Alain Royer1,2, Danielle De Sève3, Alexandre Roy1,2,4,a, and Emmanuel Cosme5

1 CARTEL, Université de Sherbrooke, Québec, Canada
2 Centre d’Études Nordiques, Québec, Canada
3 IREQ, Hydro-Québec, Québec, Canada
4 Département de Géographie, Université de Montréal, Québec, Canada
5 Institut des Géosciences de l’Environnement, IGE, UGA-CNRS, Grenoble, France

a now at: Université du Québec à Trois-Rivières, Québec, Canada

Correspondence: Fanny Larue (fanny.larue@usherbrooke.ca)

Received: 26 February 2018 – Discussion started: 13 March 2018
Accepted: 1 October 2018 – Published: 5 November 2018

Abstract. Over northeastern Canada, the amount of water stored in a snowpack, estimated by its snow water equivalent (SWE) amount, is a key variable for hydrological applications. The limited number of weather stations driving snowpack models over large and remote northern areas generates great uncertainty in SWE evolution. A data assimilation (DA) scheme was developed to improve SWE estimates by updating meteorological forcing data and snowpack states with passive microwave (PMW) satellite observations and without using any surface-based data. In this DA experiment, a particle filter with a Sequential Importance Resampling algorithm (SIR) was applied and an inflation technique of the observation error matrix was developed to avoid ensemble degeneracy. Advanced Microwave Scanning Radiometer 2 (AMSR-2) brightness temperature ($T_B$) observations were assimilated into a chain of models composed of the Crocus multilayer snowpack model and radiative transfer models. The microwave snow emission model (Dense Media Radiative Transfer – Multi-Layer model, DMRT-ML), the vegetation transmissivity model ($\omega$-$\tau_{opt}$), and atmospheric and soil radiative transfer models were calibrated to simulate the contributions from the snowpack, the vegetation, and the soil, respectively, at the top of the atmosphere. DA experiments were performed for 12 stations where daily continuous SWE measurements were acquired over 4 winters (2012–2016). Best SWE estimates are obtained with the assimilation of the $T_B$ at 11, 19, and 37 GHz in vertical polarizations. The overall SWE bias is reduced by 68 % compared to the original SWE simulations, from 23.7 kg m$^{-2}$ without assimilation to 7.5 kg m$^{-2}$ with the assimilation of the three frequencies. The overall SWE relative percentage of error (RPE) is 14.1 % (19 % without assimilation) for sites with a fraction of forest cover below 75 %, which is in the range of accuracy needed for hydrological applications. This research opens the way for global applications to improve SWE estimates over large and remote areas, even when vegetation contributions are up to 50 % of the PMW signal.

1 Introduction

In Quebec, eastern Canada, snowmelt runoff has become a major economic issue and plays a considerable role in flood events (Perry, 2000). Good forecasting of this water supply is essential in optimizing hydroelectric dam management. The amount of water stored in a snowpack is estimated by the snow water equivalent (SWE). Accurately predicting the evolution of the SWE is challenging over large and remote areas due to the high spatial and temporal variability of the snowpack and to the lack of in situ data, which are time-consuming and expensive to measure. Current operational hydrological forecasting models used by Hydro-Québec, one of the larger energy producers in North America, rely on the interpolation of surface snow survey measurements (Tapsoba et al., 2005; Brown et al., 2018). It has been shown that the highest uncertainties in hydrological forecasting related to
The assimilation of satellite observations is a promising approach for reducing uncertainties related to the lack of in situ data (Pietroniro and Leconte, 2005; Durand et al., 2009; Touré et al., 2011; De Lannoy et al., 2012; DeChant and Moradkhani, 2011; Andreadis and Lettenmaier, 2012; Kwon et al., 2017). In particular, passive microwave (PMW) satellite observations, which measure brightness temperatures (\(T_B\)), are sensitive to the volume of snow and provide information at a good temporal and spatial coverage (Halikainen, 1984; Chang et al., 1996; Tedesco et al., 2004). It has been shown that the assimilation of PMW satellite data into snow models adds valuable information to compensate for initialization errors and improve SWE simulated by snow models (Sun et al., 2004). These approaches appear to be very promising to evaluate and predict water resources but are still under development for further use in operational hydrological applications (Xu et al., 2014). Larue et al. (2017) showed that the GlobSnow-2 SWE product (Takala et al., 2011), which assimilates both \(T_B\) satellite data and local snow depth observations, was not accurate enough for hydrological modeling, mainly because of its dependence on in situ data in remote areas.

The main difficulty in the assimilation of PMW satellite observations in boreal forest areas is the quantification of all the contributions that affect the measured signal. PMW satellite observations have a low spatial resolution (\(\sim 10 \times 10 \text{ km}^2\)) and satellite sensors measure many contributions in addition to the PMW emission from the volume of the snowpack (vegetation canopy, ice crust, frozen/unfrozen soil, lakes, moisture in the snow, topography, etc.) (Kelly et al., 2003; Koenig and Forster, 2004). In boreal areas, the PMW emission from the forest canopy within a pixel can contribute up to half of the PMW signal measured by satellite sensors (Roy et al., 2012, 2016). This contribution does not only depend on the fraction of forest cover, but also on the biomass (liquid water content, LWC), the vegetation volume, and the canopy structure (stem, leaf, trunk) (Franklin, 1987). To adjust snowpack model simulations, several studies suggest using radiative transfer models, coupled to a snowpack model, to take into account the different contributions to the PMW signal at the top of the atmosphere and to directly assimilate PMW satellite observations (Brucker et al., 2011; Durand et al., 2011; Langlois et al., 2012; Roy et al., 2016). However, the assimilation of PMW must be used with care, and a good understanding of the interactions between the properties and microwave emission of the snowpack is crucial to avoid degradation of the SWE estimates. For instance, the assimilation of passive microwave in wet snow conditions can introduce large uncertainties since the presence of liquid water in the snowpack increases \(T_B\), whereas increases in snow grain size decrease the brightness temperature independent of any change in SWE (Klehmet et al., 2013). The assimilation of PMW thus can help to adjust the modeled snowpack states during the winter, but it cannot be used at the beginning and at the end of the season (snowmelt periods).

This paper aims at developing and validating the assimilation of PMW satellite observations for SWE improvements over Quebec by adjusting meteorological forcing data and simulated snowpack states without using any surface-based observations. Advanced Microwave Scanning Radiometer 2 (AMSR-2) satellite sensors provide the \(T_B\) observations at 11, 19, and 37 GHz. The data assimilation scheme (DA) is a Sequential Importance Resampling particle filter (referred to as PF-SIR) (Van Leeuwen, 2009, 2014). The PMW emission from the snowpack is computed by using the Crocus snowpack model (Brun et al., 1989) coupled to a microwave snow emission model, the Dense Media Radiative Transfer – Multi-Layer model (DMRT-ML) (Picard et al., 2013). This scheme is further referred as the Crocus/DMRT-ML chain and was previously calibrated over Quebec (Larue et al., 2018). As a first step, the previous study of Larue et al. (2018) tested the feasibility of the DA scheme in a controlled environment by using synthetic \(T_B\) observations, obtained by running the Crocus/DMRT-ML chain with perturbed meteorological forcings. The results showed SWE root mean square error (RMSE) reduced by 82 % with the multivariate assimilation of \(T_B\) at 37, 19, and 11 GHz in vertical polarizations, compared to SWE RMSE without assimilation. In the present study, the same DA setup as described in Larue et al. (2018) was implemented, except that real satellite observations were used. For the assimilation of satellite data, the challenge is to accurately simulate the \(T_B\) measured at the top of the atmosphere (\(T_{B\ TOA}\)) by including contributions other than snow (i.e., soil, vegetation, and atmosphere). The vegetation transmissivity model (\(\omega\)), the Wegmüller and Mätzler (1999) soil emission model, and the Liebe (1989) atmospheric emission model were added and calibrated to simulate the PMW emission of satellite observations (Roy et al., 2016).

The specific objectives of this paper were thus to (1) calibrate the soil and the vegetation radiative transfer models coupled with the Crocus/DMRT-ML chain to simulate \(T_{B\ TOA}\) over several years (2012 to 2016); and (2) evaluate the performance of the assimilation of PMW data in Crocus using SWE measurements obtained over 12 reference nivometric stations from 2012 to 2016 (43 winters). This paper opens the way to a functional spatialized method for improving SWE estimates over large and remote areas without using surface-based data.
To evaluate SWE simulations, SWE measurements were acquired from 2012 to 2016 by 12 nivometric stations (see numbered stations on Fig. 1), located through a north–south gradient in Quebec. This SWE database (coordinates, sensors, operating period, etc.) was fully described in Larue et al. (2018). Table 1 describes the main station characteristics, including the mean maximum SWE values over operating periods. Daily SWE measurements were derived from gamma ray SWE sensors (Campbell Scientific CS725, “GMON”) with an average error of +5% (Choquette et al., 2008). Two stations (nos. 5 and 12) were located in the subarctic ecoclimatic zone (53–54°N, James Bay area), eight in the coniferous boreal zone (46–48°N), and two (Nos. 4 and 11) in a mixed forest area in southern Quebec (45.3°N). Sensors were calibrated by Hydro-Québec from numerous field measurement campaigns during the first year following their installation.

A total of 43 winters were studied (Table 1). These winters were all very different. Winter 2012–2013 had the lowest snow accumulation in 10 years (165 cm), whereas winter 2013–2014 was very snowy (379 cm) compared to the average snow accumulation (217 cm). Winter 2014–2015 was unusually cold (3° below average temperatures), and winter 2015–2016 was the warmest in 60 years (statistics can be found at http://www.mddep.gouv.qc.ca, last access: 18 October 2018).

2.2 General setup

Figure 2 shows the general methodology developed to simulate and to assimilate AMSR-2 satellite observations into the snowpack model.

To simulate the signal measured by satellite sensors at the top of the atmosphere (\(T_{B_{\text{TOA}}}\)), a chain of models was implemented and calibrated over eastern Canada. The 3-hourly continuous atmospheric forcing database provided by the Global Environmental Multiscale weather prediction model (referred to as “GEM”; Coté et al., 1998) was used to drive the multilayer Crocus snowpack model (described in Sect. 3.2.1). Each GEM grid cell has a spatial resolution of 10 × 10 km², which is on the same order as the observation scale. The Crocus model updates the snowpack every 15 min by interpolating meteorological inputs, but in this study we used daily Crocus outputs (SWE, snow depth, density, etc.) computed at 14:00 local time (19:00 UTC), in agreement with the AMSR-2 pass (Sect. 3.1.1). The DMRT-ML radiative transfer model (Sect. 3.2.1), driven with Crocus outputs, was used to simulate the PMW emission from the modeled snowpack (referred to as “\(T_{B_{\text{snow}}}\)”) at 11, 19, and 37 GHz, at vertical and horizontal polarizations (“V-pol” and “H-pol”, respectively). The contribution of the atmosphere was estimated by using an atmospheric model (Liebe, 1989) driven with the total precipitable water integrated over 28 atmospheric layers and provided by GEM (Sect. 3.3). The surface emissivity for a rough soil was deduced by calibrating the Wegmüller and Mätzler (1999) soil model, and vegetation contributions were quantified with the \((\omega, \tau_{\text{opt}})\) radiative transfer model (Sect. 3.3). To take canopy emissivity variability into account, the inversions of the \((\omega, \tau_{\text{top}})\) parameters were linked to the 4-day leaf area index (LAI) product from MODIS data (1 × 1 km²), averaged for each AMSR-2 grid cell (10 × 10 km²) (Sect. 3.3). These inversions of soil...
Figure 2. Methodological scheme describing the DA scheme in the chain of models for SWE retrievals by updating perturbed atmospheric forcing data and snowpack states (“$F_t$” and “$x_t$”, respectively; see Sect. 3.4).

Table 1. Characteristics of the nivometric SWE stations: site number, latitude (Lat.), longitude (Long.) and elevation (El., a.s.l. in meters) of stations. Dist. GEM–station is the distance between the station and the center of the associated GEM grid cell (with GEM, Global Environmental Multiscale weather prediction model; Sect. 2.2). The time period of observations, average of the maximum observed data over the studied period, and data provider are given (HQ: Hydro-Québec, U. Sherb: University of Sherbrooke, U. Laval: University of Laval).

<table>
<thead>
<tr>
<th>Sites no.</th>
<th>Lat.</th>
<th>Long.</th>
<th>El.</th>
<th>Dist. GEM–station (km)</th>
<th>Time period</th>
<th>Mean maximum SWE value (kg m$^{-2}$)</th>
<th>Data provider</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>48.3</td>
<td>-74.1</td>
<td>100</td>
<td>3.4</td>
<td>2012–2016</td>
<td>272</td>
<td>HQ</td>
</tr>
<tr>
<td>2</td>
<td>48.9</td>
<td>-74.2</td>
<td>100</td>
<td>4.9</td>
<td>2012–2016</td>
<td>277</td>
<td>HQ</td>
</tr>
<tr>
<td>3</td>
<td>47.9</td>
<td>-72.9</td>
<td>100</td>
<td>4.7</td>
<td>2012–2016</td>
<td>252</td>
<td>HQ</td>
</tr>
<tr>
<td>4</td>
<td>46.6</td>
<td>-72.8</td>
<td>136</td>
<td>4.2</td>
<td>2012–2016</td>
<td>253</td>
<td>HQ</td>
</tr>
<tr>
<td>5</td>
<td>53.7</td>
<td>-78.2</td>
<td>103</td>
<td>4.2</td>
<td>2012–2016</td>
<td>213</td>
<td>HQ</td>
</tr>
<tr>
<td>6</td>
<td>46.7</td>
<td>-76.0</td>
<td>229</td>
<td>2.3</td>
<td>2012–2016</td>
<td>161</td>
<td>HQ</td>
</tr>
<tr>
<td>7</td>
<td>47.0</td>
<td>-74.3</td>
<td>469</td>
<td>3.3</td>
<td>2012–2016</td>
<td>235</td>
<td>HQ</td>
</tr>
<tr>
<td>8</td>
<td>46.9</td>
<td>-76.4</td>
<td>330</td>
<td>1.8</td>
<td>2012–2016</td>
<td>212</td>
<td>HQ</td>
</tr>
<tr>
<td>9</td>
<td>46.9</td>
<td>-73.7</td>
<td>372</td>
<td>1.9</td>
<td>2012–2016</td>
<td>180</td>
<td>HQ</td>
</tr>
<tr>
<td>10</td>
<td>47.7</td>
<td>-73.6</td>
<td>398</td>
<td>3.5</td>
<td>2012–2016</td>
<td>202</td>
<td>HQ</td>
</tr>
<tr>
<td>11</td>
<td>47.3</td>
<td>-71.2</td>
<td>669</td>
<td>2.6</td>
<td>2015–2016</td>
<td>396</td>
<td>U. Laval</td>
</tr>
<tr>
<td>12</td>
<td>53.4</td>
<td>-75.0</td>
<td>389</td>
<td>4.0</td>
<td>2014–2016</td>
<td>211</td>
<td>U. Sherb</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2012–2016</td>
<td>237</td>
<td></td>
</tr>
</tbody>
</table>

and vegetation parameters were performed over the summer period to avoid bias due to the presence of the snowpack.

The brightness temperatures ($T_{B_s}$) measured by AMSR-2 satellite sensors were assimilated in a DA scheme (see Sect. 3.4). Raleigh et al. (2015) have shown that meteorological forcing data are the major sources of errors in snow model simulations. Hence, we assume here that the uncertainties of GEM meteorological forcing data are the only sources of errors in the $T_B$ modeling. It is very difficult to quantify modeling errors due to physical simplifications inside the model due to the spatial scale of the observations. Further studies are needed to estimate these errors over the study area and to take them into account in the DA experiment. The observation error was assumed to be known and the modeling errors were estimated by perturbing selected meteorological forcing variables. An ensemble of 150 $T_B$ simulations was obtained and the distribution of these prior estimates represents the modeling error in response to GEM uncertainties. A particle filter with an SIR algorithm was used to update the simulated $T_{B TOA}$ over the winter by adjusting meteorological forcing data and snowpack states (posterior estimates) when an observation was available (Fig. 2).

Several configurations of the DA scheme were tested over three evaluation sites representing different environmental
conditions. The best configuration was evaluated over the validation reference sites from 2012 to 2016 (for 43 winters; Sect. 3.4).

Comparing data simulated at the station against model cells involves uncertainty due to spatial variations of the snowpack and land cover. This is a well-known problem for model validation studies and we assume here that the high number of sites (12 SWE stations or 43 snowpack simulations) provides a useful assessment of simulations. It is also known that the spatial localization of measurements can lead to some biases (Molotch and Bales, 2005). To diversify its measurements, Hydro-Québec has installed two SWE sensors in the forest, and not in a clearing as is the usual practice for ease of maintenance.

3 Materials and methods

3.1 Data

3.1.1 AMSR-2 observations

AMSR-2 satellite sensors (Imaoka et al., 2010) provide PMW satellite observations on the 11 (10.7), 19, and 37 GHz channels at V-pol and H-pol. Images produced by AMSR-2 are freely available on the Japan Aerospace Exploration Agency (JAXA) website. This study used the Level 3 Version 2 product, which provides daily $T_{\text{BS}}$ normalized on a North Hemisphere polar stereographic projection with a spatial resolution of $10 \times 10 \text{ km}^2$ (see https://gportal.jaxa.jp/, last access: 18 October 2018, for the specifications of the projection), from 1 August 2012 to 1 July 2016. $T_{\text{BS}}$ from AMSR-2 are computed twice a day: around 13:30 local time, or 17:30 UTC (ascending pass), and around 01:30 local time, or 05:30 UTC (descending pass). Only the ascending pass was used in this study since the snowpack was computed once a day at 14:00 (local time). The use of the ascending pass allowed the nighttime refreezing process to be avoided. To reduce observation errors due to the daytime melting process, the approach was evaluated during the dry snow period, from December to mid-March. This aspect is further discussed in Sect. 5.1.

3.1.2 LAI MODIS data

The 4-day LAI product provided by MODIS TERRA data (MOD15A3; Myneni et al., 2002) was used to characterize the vegetation contributions to the total emissivity (Fig. 2). The product has a spatial resolution of $1 \times 1 \text{ km}^2$ and was resampled on the AMSR-2 grid of $10 \times 10 \text{ km}^2$ by averaging all LAI data within each AMSR-2 grid cell (referred to as “LAI_{AMSR-2}”). For each site, Table 2 describes the summer and winter average values (“LAI_{summer}” and “LAI_{winter}”) calculated using LAI_{AMSR-2} from 1 July to 31 August and from 1 January to 1 March over the 2012 to 2016 time period, respectively (Roy et al., 2014).

3.1.3 Land cover map of Canada

The land cover map of Canada Circa 2000 (available at http://www.geobase.ca/geobase/en/data/landcover/index.html, last access: 18 October 2018) (referred to as “LCC”) was used to extract the fraction of forest cover (“f_{cover}”) within each AMSR-2 grid cell. This product provides the percentage of coniferous, herbaceous, deciduous, and water areas with a spatial resolution of $1 \times 1 \text{ km}^2$ and was resampled to generate average values within each $10 \times 10 \text{ km}^2$ AMSR-2 grid cell. Table 2 shows the fractions of forest cover provided by the LCC and resampled over AMSR-2 grid cells for each site. As expected, Sites 5 and 12, which are located in the subarctic area (Fig. 1), have a low $f_{cover}$ (below 32%). The other sites in boreal areas have an $f_{cover}$ of up to 60%. Sites 6 and 9 are in particularly densely forested areas, with a high $f_{cover}$ (up to 80%). The measured $T_{B}$ signal can be significantly affected by the forest and the signature of the underlying snow is attenuated during the winter period in such densely forested areas. The sensitivity of the DA scheme to the $f_{cover}$ was analyzed for sites with an $f_{cover}$ above and below 75% (Sect. 4.2.1).

Moreover, the presence of lakes can affect the PMW signal. Lake ice (when snow cover is absent) increases the PMW signal at high frequencies, and at low frequencies, the contribution of water bodies acts as a reflector and the emissivity remains low (De Sève et al., 1999). With snow cover on lakes, the different snow states on the lakes compared to snow cover under forest also modified the emitted signal (see Derksen et al., 2012, 2014). Nevertheless, we made the hypothesis that these impacts were negligible over our studied sites, which have lake water fractions under 7% within their AMSR-2 grid cells (Table 2) (masks are generally applied for water fractions of up to 20%; Takala et al., 2011).

3.2 Simulation of the PMW emission from the snowpack

3.2.1 Coupling of Crocus and DMRT-ML

The chain of models developed to simulate $T_{B_{\text{snow}}}$ is identical to that of Larue et al. (2018), so only a brief description of the approach is detailed here (see Fig. 2).

The Crocus snowpack evolution model (Brun et al., 1989, 1992; Vionnet et al., 2012) is coupled with the ISBA land surface model within the SURFEX interface (Surface Exter nalisiée, in French) (Decharme et al., 2011; Masson, 2013). SURFEX/ISBA/Crocus (hereafter referred to as “Crocus”) computes the evolution of the physical properties of the snowpack and the underlying ground (soil). In particular, the snow layers are modeled with a set of variables representing the morphological properties of snow grains (shape and size), including the specific surface area (SSA), which is one of the most sensitive variables for snowpack emission simulations. The snow microstructure evolves in time according
to semi-empirical laws (Vionnet et al., 2012). Crocus is the only model able to simulate the SSA as a prognostic variable (rather than as a diagnostic variable) by using the formulations of Carmagnola et al. (2014). The number of snow layers is dynamic and evolves with physical properties updated at each time step. The maximum number of simulated snow layers was fixed at 15 in this study as a compromise between accuracy and computing time (not shown). Configuration and initialization of the Crocus snowpack model are the same as described in Larue et al. (2018).

$T_{B,\text{snow}}$ was computed by driving the radiative transfer model DMRT-ML with Crocus outputs. The DMRT-ML model is well detailed in the literature (Tsang et al., 1992; Tsang and Kong, 2001; Picard et al., 2013, Royer et al., 2017), so only the calibration is described here. Snow grain size, and more generally snow microstructure, are factors that most affect the accuracy of simulated PMW emission from a snowpack as they determine the strength of scattering mechanisms in the snowpack at high frequencies used (Roy et al., 2013; Leppänen et al., 2015; Sandells et al., 2017, Larue et al., 2018). In DMRT-ML, snow grains are represented as spheres of ice with variable interactions between them. The potential formation of clusters of grains, which increases the effective snow grain size, is not taken into account, generating uncertainties (Picard et al., 2013). Several studies have shown that DMRT-ML needed an effective scaling factor to represent the stickiness between snow grains and to correct the snow microstructure representation (Brucker et al., 2011; Roy et al., 2013; Royer et al., 2017). Larue et al. (2018) have shown that a mean snow stickiness parameter ($\tau_{\text{snow}}$) of 0.17 was optimal to simulate $T_{B,\text{snow}}$ over boreal snow in Quebec (RMSE of 27 K) when DMRT-ML is driven by Crocus snow profiles. This constant $\tau_{\text{snow}}$ value was thus used in the implemented chain of models (Sect. 3.4.3; Experiment A and B). Nevertheless, this effective parameter could change with snow type (Royer et al., 2017; Larue et al., 2018). Hence, the quality of the DA scheme with the use of the $\tau_{\text{snow}}$ parameter as a free variable was studied (Sect. 3.4.3, Experiment C).

### 3.2.2 Ice lens detection algorithm

Since ice lenses (“ILs”) within a snowpack significantly reduce $T_B$ mainly at H-pol (Montpetit et al., 2013; Roy et al., 2016), ice layers must be detected and added in the simulated Crocus snow profiles to improve $T_{B,\text{snow}}$ simulations. $T_B$ in H-pol are much more attenuated by the presence of an IL than $T_B$ in V-pol, since the coefficient of reflectivity is stronger in H-pol (Montpetit et al., 2013). Therefore, by following the daily evolution of the PMW emission from the snowpack with AMSR-2 observations, the formation of an IL can be detected by using a threshold on the polarization ratio (PR) defined by Cavaliere et al. (1984) for a given frequency ($\nu$):}

$$
\text{PR} (\nu) = \frac{T_B (\nu, \text{V-pol}) - T_B (\nu, \text{H-pol})}{T_B (\nu, \text{V-pol}) + T_B (\nu, \text{H-pol})}.  \tag{1}
$$

In this study, an IL was added on the top of the simulated snowpack if the AMSR-2 PR(11) was above 0.06 (Roy, 2014). This IL was represented as a 1 cm layer with a density of 900 kg m$^{-3}$ and with the snow grain radius set to zero (Roy et al., 2016). The difficulty is to know how to evolve this IL in the snowpack. The Crocus snowpack model has not yet been adapted to integrate the formation of ILs and evolve them in a coherent way (Quéno et al., 2016). Nevertheless, it was shown in Larue et al. (2018) (from field measurements) that an IL of 1 cm located at 4 cm from the surface of the simulated snowpack minimized the bias of DMRT-ML sim-
3.3 Simulation of the PMW emission at the top of the atmosphere

The PMW brightness temperature \( T_{B, \text{TOA}} \) emitted at the scale of the AMSR-2 product can be written for each grid cell as

\[
T_{B, \text{TOA}} = f_{\text{season}} \cdot T_{B, \text{forest}} + (1 - f_{\text{season}}) \cdot T_{B, \text{open}} + T_{B, \text{atm}^T}, \tag{2}
\]

where \( T_{B, \text{atm}^T} \) is the ascending atmospheric contribution, estimated using the Liebe (1989) model implemented in the Helsinki University of Technology (HUT) snow emission model (Pulliainen et al., 1999). The model considers radiative transfer through the atmospheric layers and provides \( T_{B, \text{atm}^T} \) values at the satellite sensor level (Liebe, 1989) according to the precipitable water integrated for all atmospheric layers provided by GEM. \( f_{\text{season}} \) is the seasonal (winter or summer) fraction of forest cover in the AMSR-2 grid cell, \( T_{B, \text{forest}} \) is the PMW emission with vegetation contributions, and \( T_{B, \text{open}} \) is the PMW emission without vegetation contributions.

The \( f_{\text{cover}} \) values provided by the LCC map were constants, whereas these fractions of forest evolve with the season. To take into account the seasonal variation of the forest cover for the winter and summer periods (defined as the time period with and without snow, respectively) and to estimate the \( f_{\text{season}} \) used in Eq. (2), \( f_{\text{cover}} \) was linked respectively to \( \text{LAI}_{\text{winter}} \) and \( \text{LAI}_{\text{summer}} \) by comparing the \( f_{\text{cover}} \) map to the two resampled maps (both resampled on the AMSR-2 projection) throughout Quebec (not shown). The seasonal fractions of \( f_{\text{cover}} \) were related to seasonal LAIs with Eqs. (3) and (4) for summer and winter, respectively:

\[
f_{\text{summer}} = 0.9 \cdot (1 - \exp(-2.7 \cdot \text{LAI}_{\text{summer}}))^{3.2} \tag{3}
\]

\[
f_{\text{winter}} = 0.9 \cdot (1 - \exp(-16.0 \cdot \text{LAI}_{\text{winter}}))^{0.3}. \tag{4}
\]

The linear correlation between the \( f_{\text{summer}} \) values estimated from the LCC and the \( f_{\text{summer}} \) values fitted to LAI data with the Eq. (3) had a coefficient correlation \( R \) equal to 0.94 and a \( p \) value below 0.01. For the LCC \( f_{\text{winter}} \) values and the \( f_{\text{winter}} \) values fitted to the LAI data (see Eq. 4), the coefficient correlation \( R \) was equal to 0.95 and the \( p \) value was below 0.01.

3.3.1 Vegetation contributions

The PMW emission from the vegetation varies with the forest characteristics, such as the biomass, the structure of the vegetation, or the liquid water content of the canopy. In this study, the vegetation contribution was modeled with the simplified radiative transfer model \( (\omega - \tau_{\text{opt}}) \) (Mo et al., 1982), in which the parameters should be estimated by fitting the simulated \( T_{B, \text{Bs}} \) with observations (Grant et al., 2008; Roy et al., 2012). The \( \omega \) is the single scattering factor of the albedo. Given the incidence angle \( \theta = 55^\circ \) of AMSR-2 satellite sensors, the optical thickness of the vegetation \( \tau_{\text{opt}} \) was a function of the forest transmissivity \( (\gamma) \) such that \( \gamma = \exp(-\tau_{\text{opt}}/\cos\theta) \). The forest transmissivity varies with the frequency \( (\nu) \) used and is further called \( \gamma_{\nu} \). At the satellite sensor, the expression of \( T_{B, \text{TOA}} \) in boreal areas was described by Eq. (2), which can be detailed with Eqs. (5) and (6) (see Roy et al., 2012):

\[
T_{B, \text{forest}} = \left[ \gamma_{\nu} \cdot e_{\text{surf}} \cdot T_{\text{surf}} + (1 - \omega) \cdot (1 - \gamma_{\nu}) \cdot T_{\text{veg}} + \gamma_{\nu} \cdot (1 - e_{\text{surf}}) \cdot (1 - \omega) \cdot (1 - \gamma_{\nu}) \cdot T_{\text{veg}} + (1 - e_{\text{surf}}) \cdot \gamma_{\nu}^2 \cdot T_{B, \text{atm}^\text{↓}} + (1 - \gamma_{\nu}) \cdot \omega \cdot T_{B, \text{atm}^\text{↑}} \right] \gamma_{\text{atm}}, \tag{5}
\]

\[
T_{B, \text{open}} = \left[ e_{\text{surf}} \cdot T_{\text{surf}} + (1 - e_{\text{surf}}) \cdot T_{B, \text{atm}^\text{↑}} \right] \gamma_{\text{atm}}. \tag{6}
\]

where \( T_{\text{surf}} \) is the surface temperature, \( e_{\text{surf}} \) is the surface emissivity under the canopy (with or without snow) for a given frequency, and \( T_{\text{veg}} \) is the temperature of the vegetation (taken as equal to the air temperature at 2 m, provided by GEM). \( T_{B, \text{atm}^\text{↓}} \) is the descending atmospheric contributions and \( \gamma_{\text{atm}} \) is the transmittance of the atmosphere. These atmospheric contributions were modeled using the Liebe (1989) model, as were the \( T_{B, \text{atm}^\text{↑}} \) values. Thus, for snow-free conditions, only forest \( (\omega, \gamma_{\nu}) \) and soil \( (e_{\text{surf}}) \) parameters were unknown and needed to be adjusted for each site by fitting the model output to the observations.

3.3.2 Soil contributions

To deduce the surface emissivity for rough soil \( (e_{\text{surf}, p} \) for a given polarization \( p \)), the Wegmüller and Mätzler (1999) soil
model was used to calculate the surface reflectivity of rough soil under the canopy \( r_{\text{surf},p} \) for a given polarization \( p \), with or without snow by using Eqs. (7) and (8):

\[
r_{\text{surf}, H} = 1 - e_{\text{surf}, H} = \Gamma_{\text{Fresnel}, H} \cdot \exp(-\sigma_s \sqrt{a \cos \theta})
\]

\[
r_{\text{surf}, V} = 1 - e_{\text{surf}, V} = r_{\text{surf}, H} \cos \theta \beta_v,
\]

where \( r_{\text{surf},p} \) mainly depends on the surface roughness and Fresnel coefficients \( \Gamma_{\text{Fresnel}, H} \). In Eq. (7), the simplified parameter \( \sigma_s = k \sigma \) was used, where \( k \) is the wave number and \( \sigma \) the standard deviation of the surface height (in meters); \( a \) is a constant parameter fixed to \(-0.1\) (Wegmüller and Mätzler, 1999). For frozen soil, parameters derived from Montpetit et al. (2018) were used (see Sect. 4.1). For thawed soil, \( \Gamma_{\text{Fresnel}, H} \) was estimated from the dielectric constant calculated with the Dobson (1985) equations according to the soil moisture and soil temperature. These variables were computed with the Crocus model, coupled to the ISBA land surface model, and extracted daily (at 14:00, as the other variables). The soil reflectivity in vertical polarization also depends on a parameter \( \beta_v \) (Montpetit et al., 2018), which describes the polarization of the signal and is frequency-dependent. Note that we will often use the “v” subscript hereafter to denote quantities that are dependent on frequency.

Hence, the soil parameter \( e_{\text{surf}} \) was linked to the set of values \( \{\sigma_s, \beta_v\} \) and mainly evolved with soil moisture and soil temperature.

### 3.3.3 Inversions of vegetation and soil parameters

The inversion of forest \( \{\omega, \gamma_v\} \) and soil \( \{\sigma_s, \beta_v\} \) parameters was carried out in summer to avoid the bias due to the presence of a snowpack. Forest parameters \( \{\omega, \gamma_v\} \) depend on the forest characteristics, such as the biomass and the structure of the canopy at each site. They also depend on LAI, which allows the season forest emission cycle to be accounted for. Using the vegetation water content equation defined by Pampaloni and Paloscia (1986), the parameter \( \gamma_v \) was related to the 4-day LAI for a given frequency \( \nu \) with the Eq. (9):

\[
\gamma_v = e^{-b \omega^a \left( \exp\left(-\frac{\nu}{T_s}\right) - 1\right) / \cos \theta},
\]

where \( a \) and \( b \) are two constants to calibrate. To reduce the number of unknown variables, Eq. (9) was simplified to use only one constant \( \eta_v \) such as \( e_v = e^{-b \omega^a} \).

The vegetation and soil parameters were inverted by minimizing the difference between \( T_{\text{B TOA}} \) simulations and \( T_{\text{B TOA}} \) measured with AMSR-2 sensors at 11, 19, and 37 GHz in vertical polarizations. We used the same approach developed by Roy et al. (2014) since it was well adapted for PMW emission in boreal areas: the two frequency-dependent parameters \( \{\eta_v, \beta_v\} \) and two frequency-invariant parameters \( \{\omega, \sigma_s\} \) were inverted with a two-stage calibration by permuting all possible combinations of the two frequency-invariant parameters. Specifically, \( \omega \) values varied from 0.02 to 0.16 in steps of 0.01, and \( \sigma_s \) varied from 0.01 to 1.1 in steps of 0.05. This yields a total of 300 possible combinations of the frequency-invariant parameters. Then, for each possible combination of the frequency-invariant parameters, a calibration of the frequency-dependent parameters, \( \eta_v \) and \( \beta_v \), was performed for each frequency. A total of 900 frequency-dependent calibrations were thus computed. Finally, for each possible combination of the frequency-invariant parameters, the total post-calibration \( T_B \) RMSE across all three frequencies was computed. The combination of frequency-invariant parameters resulting in the lowest \( T_B \) RMSE was chosen.

\( T_{\text{B TOA}} \) were simulated from 2012 to 2016. The inversion was not very sensitive to \( \sigma_s \) (not shown) and Fig. 3 shows the optimal overall \( T_{\text{B TOA}} \) RMSE between simulated and measured \( T_{\text{B TOA}} \) for the 12 sites and for the summer period according to \( \omega \) values. Over the summer period, a \( \omega \) value at 0.07 and a \( \sigma_s \) value at 0.2 cm gave the best results but \( T_B \) RMSE is not very sensitive to this variable. The parameters \( \beta_v \) and \( \eta_v \) were optimized for each \( (\omega, \sigma_s) \) couple according to the frequency used.

### 3.4 Data assimilation setup

The DA setup is the same as the one developed in Larue et al. (2018) except that we added an inflation technique of the covariance matrix of observation errors (R matrix) to avoid ensemble degeneracy, i.e., when an ensemble collapses to a unique particle (Arunampalam et al., 2002).
3.4.1 DA framework

The DA scheme is a particle filter with a Sequential Importance Resampling algorithm (PF-SIR) that is well documented in Van Leeuwen (2009, 2014) and Gordon et al. (1993) and relatively easy to implement with a snowpack model (Dechant and Moradkhani, 2011; De Lannoy et al., 2012; Charrois et al., Larue et al., 2018). The PF-SIR represents the probability density function (pdf) of the model state with an ensemble of states (called particles), which is updated when an observation is available. An ensemble approach was preferred because of the nonlinearity of the system. Moreover, the particle filter approach can cope with the variable number of state variables resulting from the changing number of snow layers in Crocus. The created ensemble represents uncertainty in SWE and in 

Assuming that the meteorological forcing data were the only source of uncertainties, the ensemble of 

The daily ensemble of meteorological forcing data was created by perturbing selected GEM data (air temperature, wind speed, precipitation, and short- and long-wave radiation) with Gaussian noise according to their respective uncertainties (estimated in Larue et al., 2018). Meteorological forcing perturbations were propagated in time following a first-order autoregressive process to simulate their realistic temporal variations (Charrois et al., 2016). Precipitation, wind speed, and short-wave radiation (“SW_{down}”) were perturbed by a multiplicative factor centered at 1. Perturbation boundaries were fixed at −0.9 and 0.9. The air temperature was perturbed by an additive factor, with boundaries fixed at −3 and +3 K. Perturbed long-wave radiation (“LW_{down}”) was estimated with perturbed 

where 

The observation error standard deviation associated with AMSR-2 observations was assumed to be 2 K (Durand and Margulis, 2006, 2007). Note that in reality it was probably larger since it represents all mismatches between observations and simulations obtained if the model was run with “correct” inputs. This observation error cannot be easily estimated (low spatial resolution, representativeness, etc.), but it is only a sort of initial value here, since we used a covariance inflation to adjust it.

DA experiments were applied between 1 November and 1 May. To avoid wet snow conditions, which increase the emissivity of the snowpack, whereas the SWE does not change, the DA was not performed when liquid water content was observed in the modeled snowpack. This variable was computed by Crocus, driven with original meteorological forcing data. SWE values were evaluated over both the dry snow period (from 1 December to 15 March) and the whole winter (when a snowpack was detected).
simulations over the rest of the season. The other side of the coin is that a “good” observation can be ruled out if the model is not able to reproduce it, thereby reducing the accuracy of the snowpack estimation.

### 3.4.3 Experimental setup

To study the sensitivity and the quality of $T_B$ assimilation for SWE improvements, three experiments were performed.

(a) Experiment A: to test the feasibility of the DA scheme for several environmental conditions, and to find the best DA configuration to apply. $T_B$ assimilation for three representative sites was performed in a preliminary experiment for winter 2014–2015. Following a north–south gradient, we selected Site 12 ($f_{cover} = 24.2 \%$, northern coniferous area), Site 1 ($f_{cover} = 63.7 \%$, coniferous area), and Site 9 ($f_{cover} = 84.0 \%$, mixed forest area), each representing a different environmental condition. Over these three sites, we estimated the quality of the DA scheme according to the assimilated frequencies: (a) assimilation of the $T_B$ difference between 19 and 37 GHz (referred to as $\Delta T_B$ 19-37); (b) assimilation of the $\Delta T_B$ 19-37 and the $T_B$ difference between 11 and 19 GHz, in V-pol (referred to as “$\Delta T_B$ 11-19”); and (c) assimilation of the three $T_B$s at 11, 19, and 37 GHz in V-pol ($T_B^{11}, T_B^{19}, T_B^{37}$). Table 3 summarizes the experiment setup information (acronyms of the experiments, sites, time period). We used V-pol $T_B$ because H-pol $T_B$ is more sensitive to the stratigraphy of the snowpack and to the presence of liquid water (Mätzler, 1987). While the DA of $T_B$s at 11, 19, and 37 GHz in V-pol should give best results since this combination of frequencies imposes more constraints, the risk of encountering a degeneracy problem is higher. The combination of both $\Delta T_B$ 19-37 and $\Delta T_B$ 11-19 is commonly used in the literature for SWE retrievals (Chang et al., 1987; Tedesco et al., 2004; Tedesco and Nervekar, 2010). The assimilation of the $\Delta T_B$ 19-37 only was also studied to analyze the sensitivity of $T_B$ assimilation for deep snowpack when $T_B^{37}$ saturates for a SWE up to about 150 mm (Mätzler et al., 1994) and to evaluate the supply of information from 11 GHz in the assimilation of both $\Delta T_B$ 19-37 and $\Delta T_B$ 11-19 for SWE improvements.

To quantify the performance of the DA scheme, the daily RMSEs of ensembles of simulated SWE obtained with and without the DA scheme were compared by using Eq. (11):

$$RMSE_t = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (X_{sim,t,i} - X_{Obs,t})^2}, \quad (11)$$

where $N$ is the ensemble size, $X_{sim,t,i}$ is the simulated variable from the member $i$ at time $t$, and $X_{Obs,t}$ is the diagnostic variable at time $t$ obtained from AMSR-2 observations.

(b) Experiment B: the best configuration of the DA scheme (DA of the three $T_B$s at 11, 19, and 37 GHz in V-pol, called the “DA_b_TB_11, 19, 37” experiment in Table 3) was applied over the 43 winters. To estimate the accuracy for hydrological applications, the median of the resampled SWE ensemble obtained with the DA_b_TB_11, 19, 37 experiment (called “SWE_{DA}” further) was compared to SWE measurements. The median was used instead of the mean to reduce the potential impact of extreme perturbations. The evaluation of this experiment was performed by comparing SWE_{DA} RMSE and the relative percentage of error (“RPE”) values to the original SWE simulations (SWE_{Crocus}), obtained by driving Crocus with original meteorological forcing data. The RPE is defined as

$$RPE = 100 \cdot \frac{|\text{Bias}|}{\text{MEAN}_{obs}}, \quad (12)$$

The mean biases of SWE estimates obtained without and with assimilation were also compared. Performance was estimated for SWE higher than 48 kg m$^{-2}$ (about 20 cm of snow...
Table 4. Effective parameters calibrated for the 12 studied sites to quantify soil contributions $\epsilon_{\text{surf}}$ (calibrated surface roughness “cal. $\sigma_s$”) and calibrated polarization ratio (cal. $\beta_\nu$) and vegetation contributions (controlled by the calibrated (\omega, \eta_\nu) parameters “cal. $\omega$” and “cal. \eta_\nu,” according to the daily LAI) measured at the top of the atmosphere. The parameterization of frozen ground was estimated by Montpetit et al. (2018). $\epsilon_{\text{eff}}$ is the effective dielectric constant estimated with the permittivity of frozen and unfrozen soils derived from Dobson’s equations (1985). Annual and seasonal $T_{\text{B TOA}}$ RMSEs estimated for the summer and the winter period ($\text{RMSE}_{\text{summer}}$ and $\text{RMSE}_{\text{winter}}$) are calculated from 2012 to 2016 with the calibrated parameters.

<table>
<thead>
<tr>
<th>Frequency (GHz)</th>
<th>Frozen soil</th>
<th>Unfrozen soil</th>
<th>Cal. $\omega$</th>
<th>Cal. $\eta_\nu$</th>
<th>Mean $\text{RMSE}_{\text{summer}}$ (K)</th>
<th>Mean annual $\text{RMSE}$ (K)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\epsilon_{\text{eff}}$</td>
<td>$\sigma_s$ (cm)</td>
<td>$\beta_\nu$</td>
<td>$\sigma_s$ (cm)</td>
<td>$\beta_\nu$</td>
<td>$\nu$</td>
</tr>
<tr>
<td>11</td>
<td>3.18-0.006134i</td>
<td>1.08</td>
<td>0.69</td>
<td>0.01</td>
<td>8.6</td>
<td>8.4</td>
</tr>
<tr>
<td>19</td>
<td>3.42-0.00508i</td>
<td>0.19</td>
<td>0.72</td>
<td>0.07</td>
<td>0.05</td>
<td>9.1</td>
</tr>
<tr>
<td>37</td>
<td>4.47-0.32643i</td>
<td>0.42</td>
<td>0.67</td>
<td>0.23</td>
<td>10.1</td>
<td>26.0</td>
</tr>
</tbody>
</table>

SD were evident for the forest land cover type (about 5% with DMRT-ML). In Experiment C, the DA scheme was thus tested using $\omega$ and $\tau_{\text{snow}}$ as free variables in the assimilation process (called the “DA_c_TB_11,19,37” experiment). The DA_c_TB_11,19,37 experiment is identical to the DA_b_TB_11,19,37 experiment (over the 43 winters): only the state variables were changed. The $\omega$ parameter was perturbed with Gaussian noise, centered on 0.07 (calibrated), with a standard deviation of 0.02 and bounded by 0.05 and 0.12 (reasonable range of $T_{\text{B TOA}}$ RMSE values; Fig. 3). The snow stickiness parameter was perturbed by Gaussian noise, centered on 0.17, with a standard deviation of 0.15 and bounded by 0.1 and 0.46. These limits correspond to the range of $\tau_{\text{snow}}$ values extracted from Larue et al. (2018) over the same study area. The ensemble size was kept to 150 members.

4 Results

4.1 Results of model inversions

The mean optimal values of the $\eta_\nu$ and $\beta_\nu$ factors were estimated for the optimal (\omega, $\sigma_s$) set of values (0.07 and 0.2, respectively; see Table 4). The (\sigma_s, $\beta_\nu$) soil parameters are given in Table 4 and are used to estimate the $T_{\text{B TOA}}$ RMSE obtained with the calibrated chain of models. Without parameter inversions, the annual mean RMSE of the original $T_B$ simulations varies from 12.9 to 47.1 K for the three frequencies (not shown). With parameter inversions over the summer period, the three frequencies have a similar $T_B$ RMSE of 8.6–10.1 K (Table 4), while over the year (using parameters inverted over the summer period) the annual $T_{\text{B TOA}}$ RMSE significantly increases at 37 GHz due to the presence of the snowpack (26.0 K). The inversions make it possible to reduce the annual $T_{\text{B,37}}$ RMSE by 21.1 K. Figure 4a, b, and c show the multiyear $T_{\text{B TOA}}$ variations for Sites 12, 1, and 9, respectively, from 2012 to 2016 and at 37 GHz. At this
frequency, the simulated $T_{B\ TOA}$ is strongly underestimated when a snowpack is observed. This is likely due to an overestimation of the SWE or snow grain sizes since $T_{B\ 37}$ are attenuated in the snowpack as snow grains act as diffusers while the $T_{B\ 19}$ and $T_{B\ 11}$ are relatively unaffected by snow grains (RMSE_{summer} similar to RMSE_{winter} at 11 and 19 GHz; Table 2). Simulated SWE values were overestimated by 16 % and 20.2 % compared to SWE measurements for Sites 1 and 9, respectively, for winter 2014–2015. The objective of T_B assimilation is to reduce these overestimations. Note that the SWE simulated at Site 12 is underestimated by 19 %. The underestimation of $T_{B\ 37}$ can also be caused by an underestimation of the vegetation contributions. This aspect is further discussed in Sect. 5.2.

By integrating ILs within the snowpack when the PR11 is above 0.06, the annual $T_{B\ TOA}$ RMSE at 37 GHz is reduced and goes from 28.9 to 26.0 K.

In winter, the overall $T_{B\ TOA}$ RMSE (all frequencies) is equal to 17.4 K from 2012 to 2016 (not shown), similar to the overall RMSE estimated for the $T_{snow}$-calibrated DMRT-ML driven by in situ measurements in an open area and equal to 19.9 K compared to surface-based radiometric measurements in Quebec (Larue et al., 2018).

### 4.2 Results of AMSR-2 data assimilation (DA)

#### 4.2.1 Experiment A

Figure 5 shows variations of the daily RMSE of the SWE ensemble (see Eq. 11) obtained without and with DA (prior and posterior estimates) according to the combination of frequencies used as observations (DA1_DTB19-37, DA2_DTB19-37, DTB11-19 and DA3_TB_11,19,37 experiments; see Table 3). Table 5 summarizes these averaged RMSEs for the studied period (dry snow period and whole winter) for tested sites.

<table>
<thead>
<tr>
<th>Experiments</th>
<th>SWE ensemble RMSE (kg m⁻²)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. 1</td>
</tr>
<tr>
<td>RMSE_{dry-snow} (kg m⁻²)</td>
<td></td>
</tr>
<tr>
<td>Without assimilation (prior estimates)</td>
<td>50.7</td>
</tr>
<tr>
<td>DA1_DTB19-37</td>
<td>21.1</td>
</tr>
<tr>
<td>DA2_DTB19-37, DTB11-19</td>
<td>14.0</td>
</tr>
<tr>
<td>DA3_TB_11,19,37</td>
<td>10.5</td>
</tr>
<tr>
<td>RMSE_{annual} (kg m⁻²)</td>
<td></td>
</tr>
<tr>
<td>Without assimilation (prior estimates)</td>
<td>50.7</td>
</tr>
<tr>
<td>DA1_DTB19-37</td>
<td>21.1</td>
</tr>
<tr>
<td>DA2_DTB19-37, DTB11-19</td>
<td>14.0</td>
</tr>
<tr>
<td>DA3_TB_11,19,37</td>
<td>10.5</td>
</tr>
</tbody>
</table>

Over the three sites and for the dry snow period, the DA reduced the overall SWE ensemble RMSE by 43.9 %, 45.8 %, and 59.7 % with the DA1_DTB19-37, DA2_DTB19-37,DTB11-19, and DA3_TB_11,19,37 experiments, respectively, compared to the SWE ensemble RMSE obtained with prior estimates (Table 5). The assimilation of the three frequencies (DA3_TB_11,19,37) helps to improve SWE simulations, giving the lowest RMSE compared to other scenarios. The same trend is observed over the whole winter and the assimilation of the three frequencies reduces the overall SWE ensemble RMSE by 47.0 % (SWE ensemble RMSE of 22.1 kg m⁻²) compared to the SWE ensemble RMSE of prior estimates (SWE ensemble RMSE of 41.7 kg m⁻²).

In our previous work (Larue et al., 2018), we have shown a reduction of 82 % of the SWE RMSE by assimilating both the $\Delta T_{B\ 19-37}$ and $\Delta T_{B\ 11-19}$ and using synthetic observation data over a dry snow period. The differences between results using synthetic and real data in DA experiments are likely due to two aspects. Firstly, the snow model does not resolve the intra-pixel surface variability. We assumed homogeneous snow cover within the pixel in open areas, thus with no interactions between snow and vegetation. Even if we compare simulations with surface-based measurements in open areas, this could introduce large uncertainties (Roy et al., 2016). Secondly, the land cover variability and heterogeneity within each pixel also induce uncertainties in the mean $T_{B}$ simulation over a pixel ($T_{B}$ weighted by the fraction of forest cover; see Eq. 2).

Figure 6 illustrates the comparison between SWE measurements, the original SWE Crocus simulations (SWE\textsubscript{Crocus}), and the median of the SWE ensemble obtained with the DA3_TB_11,19,37 experiment. The yellow envelope corresponds to the SWE ensemble obtained without DA (prior estimates) and shows a large ensemble spread in response to meteorological forcing uncertainties. The gray envelope is the resampled SWE ensemble (posterior estimates). SWE simulations are very sensitive to the uncertainties of
Figure 4. Multi-year variations of simulated $T_B$ TOA (red dotted lines) and measured $T_B$ TOA (black full lines) from 2012 to 2016 at 37 GHz in vertical polarization: (a) Site 12 ($f_{\text{cover}}$ of 24%), (b) Site 1 ($f_{\text{cover}}$ of 64%), and (c) Site 9 ($f_{\text{cover}}$ of 84%) (see Table 2).

Figure 5. Variations of the SWE ensemble RMSE (Eq. 11) obtained with and without DA for the dry snow period (from 1 December to 15 March). Experiments are performed for (a) Site 12, (b) Site 1, and (c) Site 9 over the winter 2014–2015. The red line is the SWE ensemble RMSE obtained without DA (open loop runs), the blue line is the RMSE obtained with the DA1_TB19-37 experiment, the green dashed line the RMSE with the DA2_TB19-37,TB11-19 experiment, and the black dotted line the RMSE with the DA3_TB_11,19,37 experiment.
meteorological forcing data at the beginning of the winter season. If an event (melting or precipitation) is missed, a constant bias on SWE estimates is kept throughout the winter. For Sites 1 and 9, the DA scheme allows the correction of these uncertainties at the beginning of the season: the SWE 

Table 6. Averaged SWE RMSE, bias, and RPE (Eq. 12) over the 43 winters for original SWE simulation (SWE_{Crocus}) and assimilated SWE_{DA} (Experiment B). Statistical performances were estimated for SWE_{obs} > 48 kg m\(^{-2}\) (snow depth higher than ~ 20 cm). SWE_{obs} and SWE_{sim} are the averaged observed and simulated SWE, respectively.

<table>
<thead>
<tr>
<th></th>
<th>SWEobs (kg m(^{-2}))</th>
<th>RMSE (kg m(^{-2}))</th>
<th>Bias (kg m(^{-2}))</th>
<th>RPE %</th>
<th>SWE_{sim} (kg m(^{-2}))</th>
<th>RMSE (kg m(^{-2}))</th>
<th>Bias (kg m(^{-2}))</th>
<th>RPE %</th>
<th>SWE_{sim} (kg m(^{-2}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(f_{\text{cover}} &lt; 75%)</td>
<td>162.2</td>
<td>42.5</td>
<td>17.3</td>
<td>19.0</td>
<td>179.4</td>
<td>37.1</td>
<td>-1.2</td>
<td>14.1</td>
<td>161.0</td>
</tr>
<tr>
<td>(f_{\text{cover}} &gt; 75%)</td>
<td>139.0</td>
<td>62.0</td>
<td>47.8</td>
<td>33.9</td>
<td>186.8</td>
<td>68.0</td>
<td>40.2</td>
<td>38.0</td>
<td>179.2</td>
</tr>
<tr>
<td>Mean</td>
<td>157.3</td>
<td>45.0</td>
<td>23.7</td>
<td>22.1</td>
<td>41.2</td>
<td>7.5</td>
<td>18.4</td>
<td>164.8</td>
<td></td>
</tr>
</tbody>
</table>

4.2.2 Experiment B

The median of the resampled ensemble of SWE obtained with the DA of the three frequencies (SWE_{DA}) is used to estimate the global performance of the DA scheme for SWE improvements. Table 6 details the statistical performance of simulated SWE_{DA} compared to measurements and to the original SWE Crocus simulations (SWE_{Crocus}) over the 43 winters. Figure 7 compares the SWE_{DA}, SWE_{Crocus}, and SWE measurements (SWE_{obs}) from 2012 to 2016 for four sites with different \(f_{\text{cover}}\) taken as an example: Site 5 (\(f_{\text{cover}} = 31.5\%\)), 10 (\(f_{\text{cover}} = 61.8\%\)), 1 (\(f_{\text{cover}} = 63.7\%\)), and 9 (\(f_{\text{cover}} = 84.0\%\)). In this section, we first analyze the overall SWE improvements obtained with \(T_{B}\) assimilation compared to original SWE simulations, and the impact of the vegetation on the quality of the DA scheme is discussed.

Overall SWE improvements compared to original Crocus simulations

The overall SWE_{Crocus} RMSE, bias, and RPE are of 45.0 kg m\(^{-2}\), 23.7 kg m\(^{-2}\), and 22.1 %, respectively (Table 6). In comparison, the overall SWE_{DA} RMSE, bias, and RPE are improved and equal to 41.2 kg m\(^{-2}\), 7.5 kg m\(^{-2}\), and 18.4 %, respectively. The overall bias is reduced by 16.2 kg m\(^{-2}\) (68 % of SWE_{Crocus} bias) with the DA scheme. The DA of the three frequencies thus helps to improve SWE estimates over Quebec. Correlation between SWE_{DA} simulations and SWE measurements gives a similar \(R\) coefficient to the one obtained with SWE_{Crocus} simulations (\(R = 0.79\) and \(R = 0.78\), respectively), but the offset is significantly reduced with SWE_{DA} compared to SWE_{Crocus} (offset = 10 and 29 kg m\(^{-2}\), respectively). We analyzed the number of cases with significant improvements for the total of 43 simulations studied by considering a 5 % threshold on the bias and RMSE differences before and after assimilation. The SWE_{DA} bias is significantly reduced for 26 winters (60 % of cases) compared to original SWE simulations. However, the RMSE is significantly improved for only 35 % of simulations, and in 35 % of cases, RMSEs are similar.
Figure 6. Evolution of SWE measurements (black points) and SWE simulations. The SWE_{Crocus} is the red line, and the SWE obtained with the DA3 TB 11,19,37 experiment is the gray dotted line. The yellow envelope is the spread of the SWE ensemble obtained with open loop runs (prior estimates). The gray envelope is the spread of the SWE ensemble obtained with the assimilation of the three frequencies (posterior estimates). Both spreads are delimited by the 5th and the 95th percentiles. Experiments are computed for (a) Site 12, (b) Site 1, and (c) Site 9, over the winter 2014–2015.

Figure 7. Evolution of SWE measurements (black points), original SWE simulations (red full line), and the median of the SWE ensemble obtained with the DA_b TB 11,19,37 experiment (SWE_{DA}) (blue dotted line). The gray envelope is the spread of the SWE_{DA} ensemble (posterior estimates). Experiments are computed for (a) Site 5 ($f_{cover} = 31.5\%$), (b) Site 1 ($f_{cover} = 63.7\%$), (c) Site 9 ($f_{cover} = 84\%$), and (d) Site 10 ($f_{cover} = 61.8\%$), from 2012 to 2016.
Evaluation of $SWE_{\text{max}}$

The mean observed $SWE_{\text{max}}$ is equal to 235.6 kg m$^{-2}$ from 2012 to 2016, and the mean simulated $SWE_{\text{max}}$ is equal to 278.3 and 266.8 kg m$^{-2}$ without and with the assimilation, respectively. Compared to original SWE simulations, the DA scheme improves 63 % of $SWE_{\text{max}}$ simulations with an overall improvement of 12.2 kg m$^{-2}$, corresponding to 8 % of SWE measurements (Table 6). Such an uncertainty extended over the whole territory could have a strong impact, considering that 1 mm of SWE in the LG watershed could represent USD 1 million in hydroelectric power production (Brown and Tapsoba, 2007).

SWE accuracy for sites according to the $f_{\text{cover}}$

The overall RPE obtained with the DA scheme is below 15 % (RPE = 14.1 %) for sites with an $f_{\text{cover}}$ below 75 % (Table 6), which is the accuracy required for hydrological applications (Larue et al., 2017). Hence, the accuracy of $SWE_{\text{DA}}$ estimates, obtained without the use of any surface-based data, is very encouraging for hydrological needs in remote areas. In comparison, the GlobSnow-2 SWE product (Takala et al., 2011), which assimilates both $T_B$ and in-situ snow depth measurements, has a SWE RPE equal to 35.9 % over the same area in Quebec (Larue et al., 2017), twice the uncertainty of $SWE_{\text{DA}}$. Figure 7a and b (Sites 5 and 1) show that for a single site, the original SWE$_{\text{Crocus}}$ simulation works well for some years but can be underestimated or overestimated over other years. The DA scheme allows a more stable solution when the overall $f_{\text{cover}}$ is under 75 % (not the case for Site 9, for example).

Nevertheless, even if the overall RMSE is improved, the DA scheme does not help to improve SWE estimates for sites with an $f_{\text{cover}}$ above 75 % (RMSE of 66 kg m$^{-2}$) compared to original SWE simulations (RMSE of 62.0 kg m$^{-2}$). The presence of vegetation is a major source of uncertainty in $T_B$ atTOA simulations. The emission of the trees is superimposed on the signal emitted by the underlying snowpack and increases the $T_B$ measured at the satellite level (Chang et al., 1996; Brown et al., 2003). At same time, the canopy also attenuates the surface emission toward the satellite. These contributions are complex to quantify since it depends not only on the tree fraction within the pixel but also on the tree species and states which emit/attenuate a different PMW signal depending on their biomass (liquid water content), volume, and structure (stem, leaf, trunk) (Franklin, 1987). Also, the presence of trees modifies snow accumulation on the ground, depending on interception, shade, and sublimation effects (Dutra et al., 2011; Wang et al., 2009), which increases the spatial variability of the snowpack within the same pixel. These interactions between the vegetation and the snowpack are not taken into account in Crocus, and this might induce uncertainties due to model errors. Note that SWE sensors are mostly installed in clearings, which reduces this impact in comparisons against surface-based measurements.

Kwon et al. (2016) used a similar snow radiance assimilation system to correct SD by updating the Community Land Model, version 4 (CLM4), snow and soil states, and radiative transfer model with the assimilation of the 19 and 37 GHz of AMSR-E. Over North America, it produced significant improvements of SD for the tundra type, but also produced degradation for taiga snow class and forest land cover (7.1 % and 7.3 % degradations, respectively). In the present study, the use of a multilayer snowpack model makes it possible to represent PMW emission from the snowpack with DMRT-ML well, and to improve overall snowpack simulations with $T_B$ assimilation in boreal areas when the $f_{\text{cover}}$ is below 75 %. Kwon et al. (2017) obtained better results for areas with a high $f_{\text{cover}}$ in comparison to their previous study (Kwon et al., 2016) over North America by using the vegetation parameter $\omega$ as a free variable in the DA scheme, instead of pre-calibrating it as we chose to do. This aspect is further studied with the experiment C.

### 4.2.3 Experiment C

Table 7 shows the statistical SWE performances obtained with the DA$_{\text{c TB}}$11,19,37 experiment (see Table 3 for definitions), where $\omega$ and $\tau_{\text{snow}}$ are taken as free variables in the DA scheme (“SWE$_{\text{DA}, \omega, \tau_{\text{snow}}}””) over the 43 winters.

The overall $SWE_{\text{DA}, \omega, \tau_{\text{snow}}}$ RMSE, bias, and RPE are equal to 45.5 kg m$^{-2}$, $-13.2$ kg m$^{-2}$, and 20.7 %, respectively, very close to the statistical performances of the original SWE$_{\text{Crocus}}$ simulations. The use of $\omega$ and $\tau_{\text{snow}}$ as free variables does not help to improve SWE$_{\text{Crocus}}$ simulations for sites with an $f_{\text{cover}}$ below 75 %, but the bias is significantly reduced for sites with an $f_{\text{cover}}$ above 75 % ($-7.1$ kg m$^{-2}$ and a RPE of 17.5 %). In addition, the simulated $SWE_{\text{max}}$ is improved for 86 % of the 43 simulations (37 cases), with a reduction of the $SWE_{\text{max}}$ bias of 26.6 kg m$^{-2}$ (17 % of SWE measurements) compared to the original SWE$_{\text{Crocus}}$ simulation.

The use of pre-calibrated parameters is justified because the parameters $\omega$ and $\tau_{\text{snow}}$ were not measurable and could not be directly validated. Furthermore, if parameters are

### Table 7. Same as Table 6 but using the forest parameter $\omega$ and the snow stickiness parameter ($\tau_{\text{snow}}$) as free variables in the DA scheme (Experiment C) to improve SWE retrievals ($\text{SWE}_{\text{DA, } \omega, \tau_{\text{snow}}}$).

<table>
<thead>
<tr>
<th>$\text{SWE}_{\text{obs}}$</th>
<th>$\text{SWE}<em>{\text{DA, } \omega, \tau</em>{\text{snow}}}$ with the DA of the three frequencies</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_{\text{cover}}$&lt;75 %</td>
<td>RMSE (kg m$^{-2}$)</td>
</tr>
<tr>
<td>162.2</td>
<td>45.6</td>
</tr>
<tr>
<td>$f_{\text{cover}}$&gt;75 %</td>
<td>139.0</td>
</tr>
<tr>
<td>Mean</td>
<td>157.3</td>
</tr>
</tbody>
</table>
added as state variables in the DA scheme, a larger ensemble size in the DA scheme would be needed to improve the representativeness of $T_B$ uncertainties and to ensure the solution’s stability (or at least to prevent a degeneracy problem). The ensemble size was kept to 150 here but this DA$_c$$_{TB, 11, 19, 37}$ experiment should produce improved results with a larger ensemble size. However, this would require a significant computational effort. This study is a preliminary step of a PMW DA implementation for operational hydrological applications, so there was a need to limit computing time. These results suggest that the developed approach using pre-calibrated $\omega$ and $\tau_{\text{snow}}$ parameters helps to improve the retrievals for sites with an $f_{\text{cover}}$ below 75%, and the use of $\omega$ and $\tau_{\text{snow}}$ parameters as free variables in the DA scheme should be investigated in further work for sites with more than 75% forest cover.

5 Discussion

In this section, we discuss (a) the sensitivity of wet snow conditions for $T_B$ assimilation, and (b) the percentage of surface, vegetation, and atmospheric contributions in the PMW signal measured by satellite sensors.

5.1 Wet snow conditions

In wet snow conditions, water droplets act as emission sources (especially at frequencies above 30 GHz), and the snowpack becomes close to a blackbody (Brucker et al., 2011; Picard et al., 2013; Klehmet et al., 2013). The PMW observations are thus complex to use for SWE retrievals, especially at the end of the season before the spring snowmelt when the SWE is maximal. Figure 8 illustrates the SWE$_{\text{DA}}$ obtained with the DA of the three frequencies applied over the whole winter and when the snow is dry only (with an LWC equal to 0 kg m$^{-2}$), for Site 3 (winter 2013/2014). SWE estimates are strongly deteriorated when $T_B$ assimilation is performed in wet snow conditions. For this example, the SWE$_{\text{DA}}$ RMSE is equal to 31.1 kg m$^{-2}$ with a DA performed over the dry snow period only and significantly increases to 70.2 kg m$^{-2}$ by assimilating $T_{B_{\text{dry}}}$ over the whole winter (dry and wet snow conditions).

Here we used the LWC simulated by Crocus to detect wet and dry snow. This variable is subject to model errors and is linked to the original atmospheric forcing data and to their uncertainties. Further studies are needed to automatically detect wet snow events by using direct satellite observations. Previous studies have shown the potential to use the gradient ratio ($\text{GR} = (T_{B_{37}} - T_{B_{19}})/(T_{B_{37}} + T_{B_{19}})$) to detect rain-on-snow events in arctic areas (Langlois et al., 2017; Dolant et al., 2016), and this approach should be investigated for boreal forest areas in further work. The use of active microwave observations is also a promising approach with a good spatial resolution (Roy et al., 2010).

5.2 Land cover contributions within the simulated $T_{B_{\text{TOA}}}$

To properly assimilate PMW satellite observations, all contributions that affect the observed signal need to be well identified and quantified. The estimation of $T_{B_{\text{TOA}}}$ (see Eqs. 5 and 6) can be written as the sum of the PMW contributions of the open surface ($T_{B_{\text{surf}}}$), vegetation ($T_{B_{\text{veg}}}$), and atmosphere ($T_{B_{\text{atm}}}$) according to the fraction of forest ($f_{\text{cover}}$, estimated with the LAI as in Eqs. 2 and 3) and open area (1 $- f_{\text{cover}}$) with Eqs. (13), (14), and (15) as

$$ T_{B_{\text{surf}}} = f_{\text{cover}} \cdot [(1 - \omega) \cdot (1 - \gamma_{\nu}) \cdot T_{\text{surf}} + \gamma_{\nu} \cdot (1 - \epsilon_{\text{surf}}) \cdot (1 - \omega) \cdot (1 - \gamma_{\nu}) \cdot T_{\text{veg}}] \cdot \gamma_{\text{atm}} \tag{13} $$

$$ T_{B_{\text{veg}}} = f_{\text{cover}} \cdot [\gamma_{\nu} \cdot \epsilon_{\text{surf}} \cdot T_{\text{surf}}] \cdot \gamma_{\text{atm}} + (1 - f_{\text{cover}}) \cdot [\epsilon_{\text{surf}} \cdot T_{\text{surf}}] \cdot \gamma_{\text{atm}} \tag{14} $$

$$ T_{B_{\text{atm}}} = f_{\text{cover}} \cdot \left( (1 - \epsilon_{\text{surf}}) \cdot \gamma_{\nu}^2 \cdot T_{B_{\text{atm}}\dagger} + (1 - f_{\text{cover}}) \cdot \gamma_{\nu} \cdot \omega \cdot T_{B_{\text{atm}}} + \gamma_{\nu} \cdot \omega \cdot T_{B_{\text{atm}}} \right) + (1 - f_{\text{cover}}) \cdot \left( 1 - \epsilon_{\text{surf}} \right) \cdot T_{B_{\text{atm}}} \gamma_{\text{atm}} + T_{B_{\text{atm}}} \gamma_{\text{atm}} \right). \tag{15} $$

Figure 9 illustrates the percentage of each contribution before DA at 11, 19, and 37 GHz in V-pol from 2012 to 2016, for the summer and for the winter periods (defined when snowpack is detected) for Site 12 ($f_{\text{cover}}$ of 24.2%), Site 1 ($f_{\text{cover}}$ of 63.7%), and Site 9 ($f_{\text{cover}}$ of 84.0%). The percentages of each contribution are similar at 11 and 19 GHz. The contributions from the atmosphere are weak. As expected for all frequencies, the surface contributions increase for the winter period with the presence of the snowpack, while the vegetation contributions decrease as the LAI decreases, especially at 37 GHz. For Site 12, the surface contributions represent more than 80% of the PMW signal in winter when the vegetation contributions represent less than 10% of the PMW signal (same magnitude as atmosphere contributions). For Site 1, the surface and the vegetation contributions are similar in winter (40%–55%), whereas the vegetation contributions were more than 60% of the PMW signal in summer. For Site 9, the vegetation contributions remain the main contribution to the PMW signal in comparison to the surface contributions, even in winter (50%–70% of the PMW signal for 37–10 GHz). In this dense boreal forest area, the measured snowpack emission represents less than 30% of the measured signal, and SWE improvements using only $T_B$ observations are challenging. This high vegetation contribution (emission and attenuation) explains why the developed DA scheme does not succeed in significantly improving SWE estimates for these sites with an $f_{\text{cover}}$ exceeding 75%.

6 Summary and conclusion

An ensemble data assimilation (DA) scheme was implemented in a calibrated chain of models (Crocus/DMRT-ML,
Figure 8. Evolutions of measured SWE (black points) for Site 3 from 2013 to 2014, original SWE Crocus simulation (red full line), and SWE_{DA} obtained with a DA of the three frequencies applied for the entire winter (green dotted line) and when LWC = 0 only (blue full line). The simulated total liquid water content (LWC) in the snowpack (dotted gray lines) is also shown.

Figure 9. Percentage of surface (black), vegetation (dark gray), and atmosphere (light gray) contributions to the simulated PMW signal at the top of the atmosphere (before DA) at the three frequencies 11 (a), 19 (b), and 37 GHz (c). ID12, ID1, and ID9 are Site 12 (f_{cover} of 24.2 %), 1 (f_{cover} of 63.7 %), and 9 (f_{cover} of 84.0 %), respectively (see Table 2). Summer and winter periods are defined as periods when snowpack is observed or not.

soil, vegetation, and atmosphere radiative transfer models) to improve SWE estimates by updating forcing data and snowpack states with the AMSR-2 satellite observations. The developed approach does not use any surface-based data and was tested over a boreal forest area in Quebec (eastern Canada). The proposed DA scheme is a particle filter with a resampled SIR algorithm, using an inflation technique of the R matrix to avoid degeneracy problems. The multilayer snowpack model, Crocus, coupled to the surface land model ISBA, was used to simulate the evolution of the snowpack. The DMRT-ML, the (ω-τ_{opt}) model, an atmospheric model, and the Wegmüller and Mätzler (1999) radiative transfer model were pre-calibrated to simulate the PMW contributions from the snowpack, the vegetation, and the soil, respectively, at the top of the atmosphere. The DA scheme was performed over 43 winters (12 sites from 2012 to 2016; Table 1), only in the presence of dry snow. Ice lenses were detected and integrated in the snowpack by using a thresholding approach on the polarization ratio at 11 GHz. The study shows the following.

1. T_{BTOA} can be well simulated with the developed chain of models. By calibrating soil and forest parameters (ω = 0.07 and σ_{s} = 0.2 cm), the annual T_{BTOA} RMSE (all frequencies) is equal to 14.5 K from 2012 to 2016. This RMSE is similar to the overall RMSE estimated with the τ_{snow}-calibrated DMRT-ML model driven by...
in situ measurements in an open area (19.9 K compared
to surface-based radiometric measurements in Quebec;
Larue et al., 2018).

2. The assimilation of $T_{B_s}$ at 11, 19, and 37 GHz (V-pol)
improves the SWE estimates compared to the assimila-
tion of $\Delta T_{B_{19-37}}$ only (sensitive to snowpack depth)
or to the assimilation of both $\Delta T_{B_{19-37}}$ and $\Delta T_{B_{11-19}}$.
For three sites (with different $f_{\text{cover}}$), the SWE ensem-
ble RMSE of posterior estimates is reduced by 47 %
over the whole winter compared to the SWE ensemble
RMSE of prior estimates (open loop runs).

3. By using pre-calibrated $\omega$ and $\tau_{\text{snow}}$ parameters in the
DA scheme, the overall bias (for 43 winters) of the
original SWE$_{Crocus}$ simulations is significantly reduced
by assimilating $T_{B_s}$ at 11, 19, and 37 GHz (from 23.7
to 7.5 kg m$^{-2}$). The bias on SWE$_{\text{max}}$ is reduced by
12.2 kg m$^{-2}$ (8 % of SWE measurements). The over-
all RPE goes from 22.1 % to 18.4 %, which is close
to the range of accuracy needed for hydrological appli-
cations (SWE RPE <15 %). This accuracy is achieved
with the $T_{B_s}$ assimilation for sites with an $f_{\text{cover}}$ below
75 %, but the DA deteriorates SWE simulations for sites
with an $f_{\text{cover}}$ above 75 %. However, by using $\omega$ and
$\tau_{\text{snow}}$ as free variables, the DA significantly improves
SWE simulations for sites with an $f_{\text{cover}}$ above 75 %
(RPE = 17.5 %).

Even with the difficulties associated with quantifying all
the different factors that contribute to the PMW signal mea-
sured by satellite sensors in remote boreal areas (canopy, ice
crust, frozen/unfrozen ground, presence of lakes, moisture in
the snow, topography, etc.) (Kelly et al., 2003; Koenig and
Forster, 2004), and even when vegetation contributions are
50 % of the PMW signal, the implementation of a DA scheme
in a well-calibrated chain of models reduces SWE uncertain-
ties without using any surface-based data. This assimilation
scheme can be easily implemented in an operational system
using real satellite-borne observations, despite the relatively
significant computing time required. This research opens the
way for global applications to obtain more accurate SWE es-
timates over large and remote areas where few meteorologi-
cal weather stations are present.

Data availability. The daily SWE data provided by Hydro-Québec
are used for hydrological purposes and are not available to the
public due to legal constraints on the data’s availability. For the
SWE and SD data, and field campaign measurements provided by
the University of Sherbrooke, please contact the coauthor Alain
Royer (Alain.Royer@USherbrooke.ca). Meteorological GEM data
are freely available on the government of ECCC’s website (Envi-
ronment and Climate Change Canada) at https://weather.gc.ca/grib/
grib2_reg_10km_e.html, last access: 10 October 2018). Other data
used are listed in the references.
Appendix A: Online adjustment of the observation error covariance matrix $R$

Online adjustment of covariance matrices in data assimilation is quite a common approach with the ensemble Kalman filter (Dee, 1995; Miyoshi, 2011; Brankart et al., 2010) but not with the particle filter. However, in many implementations of the particle filter, the measurement pdf is considered Gaussian, so the particle weights are computed using the observation error covariance matrix $R$. This matrix can therefore also be subject to adjustment in the context of the particle filter. Online adjustment can be and is often performed by tuning a simple inflation of the initial covariance matrix. This is the approach chosen here.

Noting $\delta_i = y - h(x_i)$, the innovation for particle $i$, the weight of this particle, is

$$\text{w}_i = \frac{\text{w}_i}{\sum_j \text{w}_j}, \quad (A1)$$

where

$$\text{w}_i = \exp\left(-\frac{1}{2} \delta_i^T R^{-1} \delta_i\right). \quad (A2)$$

An inflation of matrix $R$ by a factor $1/\alpha$ (larger than 1) can be interpreted as an exponent $\alpha$ (smaller than 1) to $\text{w}_i$. Because the weights $\text{w}_i$ are nonlinear functions of $R$, inflating $R$ tends to flatten their distribution. Online adjustment consists in finding a value for $\alpha$ that flattens the distribution of weights to the point at which $N_{\text{keep}}$ particles are selected with certainty, $N_{\text{keep}}$ being a number to be prescribed. If the number $N_{\text{keep}}$ is fixed, when the resampling step is performed using Arakawa’s procedure (Arakawa, 1997), the weight of the $N_{\text{keep}}$–th particle to be selected, $\text{w}_{\text{keep}}$, must become equal to $\text{w}_{\text{ref}} = 1/N_{\text{keep}}$. Consequently,

$$\text{w}_{\text{keep}} = \left(\frac{\text{w}_{\text{keep}}}{\text{w}_j}\right)^{\alpha} = \text{w}_{\text{ref}}, \quad (A3)$$

or, written differently after taking the logarithm:

$$\alpha = \left(\log(\text{w}_{\text{ref}}) + \log\left(\sum_i (\text{w}_j)^{\alpha}\right)\right) / \log(\text{w}_{\text{keep}}). \quad (A4)$$

This equation for $\alpha$ is not solvable analytically. Instead, we find $\alpha$ after the convergence of the series:

$$\alpha_n = \left(\log(\text{w}_{\text{ref}}) + \log\left(\sum_i (\text{w}_j)^{\alpha_{n-1}}\right)\right) / \log(\text{w}_{\text{keep}}). \quad (A5)$$

The result of this adjustment is illustrated in Fig. A1. The blue dots show the first 20 weights of a sorted distribution for an ensemble of 50 particles strongly prone to degeneracy: only 4 particles have a weight larger than $1/50 = 0.02$. The minimum number of particles to be selected is fixed to $N_{\text{keep}} = 10$. After the adjustment procedure, the identified inflation factor for matrix $R$ is $3.6 (\alpha = 0.277)$, and the weight $\text{w}_{\text{keep}}$ of the 10th particle is exactly equal to 0.02.

Obviously, this procedure is only used if the number of selected particles is below the $N_{\text{keep}}$ threshold with the initial weights.
Author contributions. FL, AR, DS, and EC contributed to the conception and design of the work. All authors contributed to the acquisition, analysis, and interpretation of data. FL created new software used in the work. FL and AR wrote the manuscript, and all authors contributed to revisions of the manuscript.

Competing interests. The authors declare that they have no conflict of interest.

Special issue statement. This article is part of the special issue “Integration of Earth observations and models for global water resource assessment”. It is not associated with a conference.

Acknowledgements. The authors would like to thank the data providers: Hydro-Québec, Environment Canada (CMC-ECCC), and the University of Laval. This project was supported by financial contributions from NSERC, Canada, FRQ-NT and MITACS Quebec, and the CFQCU France–Quebec collaboration program.

Edited by: Jaap Schellekens
Reviewed by: two anonymous referees

References


Langlois, A. et al.: Detection of rain-on-snow (ROS) events and ice layer formation using passive microwave ra-
Molotch, N. P. and Bales, R. C.: Scaling snow observations

Mo, T., Choudhury, B. J., Schmugge, T. J., Wang, J. R., and Jackson,

Meissner, T. and Wentz, F.: Intercalibration of AMSR-E and

Mätzler, C.: Passive microwave signatures of landscapes in winter,

Mätzler, C.: Applications of the interaction of microwaves with

Luojus, K., Pulliainen, J., Takala, M., Lemmetyinen, J., Smolans-


Masson, V., Le Moigne, P., Martin, E., Faroux, S., Alias, A.,

Accola, R., Belamari, S., Barbu, A., Boone, A., Bouyssel, F.,

Brousseau, P., Brun, E., Calvet, J.-C., Carrer, D., Decharme, B.,

Delire, C., Donier, S., Esaoukni, K., Gibelien, A.-L., Giordani, H.,


