Impact of Assimilating Radar Solid Precipitation Data in the Canadian Precipitation Analysis System

Florence L. Beaudry, a Stephane Belair, b Julie M. Thebault, a Dikra Khedhaouiria, b Franck Lespinas, c Daniel Michelson, d Pei-Ning Feng, a and Catherine Aubry a

a Centre ESCR, Department of Earth and Atmospheric Sciences, Université du Québec à Montréal, Montréal, Québec, Canada
b Atmospheric Science and Technology, Environment and Climate Change Canada, Dorval, Québec, Canada
c Canadian Centre for Meteorological and Environmental Prediction, Environment and Climate Change Canada, Dorval, Québec, Canada
d Meteorological Research Division, Environment and Climate Change Canada, Toronto, Ontario, Canada

ABSTRACT: The Canadian Precipitation Analysis (CaPA) system provides near-real-time precipitation analyses over Canada by combining observations with short-term numerical weather prediction forecasts. CaPA’s snowfall estimates suffer from the lack of accurate solid precipitation measurements to correct the first-guess estimate. Weather radars have the potential to add precipitation measurements to CaPA in all seasons but are not assimilated in winter due to radar snowfall estimate imprecision and lack of precipitation gauges for calibration. The objective of this study is to assess the impact of assimilating Canadian dual-polarized radar-based snowfall data in CaPA to improve precipitation estimates. Two sets of experiments were conducted to evaluate the impact of including radar snowfall retrievals, one set using the high-resolution CaPA (HRDPA) with the currently operational quality control configuration and another increasing the number of assimilated surface observations by relaxing quality control. Experiments spanned two winter seasons (2021 and 2022) in central Canada, covering part of the entire CaPA domain. The results showed that the assimilation of radar-based snowfall data improved CaPA’s precipitation estimates 81.75% of the time for 0.5-mm precipitation thresholds. An increase in the probability of detection together with a decrease in the false alarm ratio suggested an improvement of the precipitation spatial distribution and estimation accuracy. Additionally, the results showed improvements for both precipitation mass and frequency biases for low precipitation amounts. For larger thresholds, the frequency bias was degraded. The results also indicated that the assimilation of dual-polarization radar data is beneficial for the two CaPA configurations tested in this study.

KEYWORDS: Atmosphere; North America; Radars/Radar observations; Numerical weather prediction/forecasting

1. Introduction

Knowledge of the amount, distribution, and intensity of precipitation is crucial since it is one of the most relevant meteorological variables for society (Balica et al. 2013; Prein et al. 2016). Real-time quantitative precipitation estimates (QPEs) are essential for mitigating floods and droughts (AghaKouchak et al. 2015; Beneyto et al. 2020) and monitoring snow coverage, water resources, and hydropower supply (Beneyto et al. 2020; Prein et al. 2016; Zarenistanak et al. 2015). Estimating precipitation represents, however, a major challenge, especially in a vast country such as Canada with many remote and sparsely observed regions. To address the need for accurate and timely precipitation information, it is critical to improve the precipitation estimation systems across the country.

The Canadian Precipitation Analysis (CaPA) system was developed and is maintained by Environment and Climate Change Canada (ECCC) since 2003 (Mahfouf et al. 2007). The system produces near-real-time precipitation analyses over North America at fine spatial resolution. The CaPA products are mostly used for hydrological modeling (Abbasnezhadi 2017; Deacu et al. 2012) and agricultural applications (NIDIS 2015), but they are also used to initialize ECCC’s prediction models to verify numerical weather forecasts and to support climate services (Fortin et al. 2018). For example, CaPA provides information to the Canadian Land Data Assimilation System (CaLDAS) used for ECCC’s operational production of soil moisture, terrestrial snow, and surface temperature analyses (Carrera et al. 2015, 2019).

CaPA provides gridded precipitation accumulation estimates over 6-h periods, ending at 0000, 0600, 1200, and 1800 UTC, and over 24-h periods, ending at 0600 and 1200 UTC, at a grid spacing of 2.5 and 10 km. The QPEs are produced by combining a first-guess field issued from a numerical weather prediction (NWP) system’s short-range forecast with observations from weather stations (Mekis et al. 2018), radars, and satellite data for liquid precipitation. Further details on CaPA’s characteristics and improvements over the years are available in Lespinas et al. (2015), Mahfouf et al. (2007), and Fortin et al. (2018). Despite the improvements made to the CaPA system over the years, the estimation of solid precipitation remains a difficult problem to solve. The background field however is of higher quality during winter compared to summer due to the type of precipitation (stratiform vs convective).

In winter, CaPA’s performance is limited by the lack of accurate precipitation measurements and by the fact that many surface observations from gauges do not pass its quality...
control when the measured 2-m temperature is less than 0°C and the wind speed at 2 m is greater than 0.6 m s⁻¹. This rejection attempts to minimize the wind-induced snowfall undercatch errors by automated weather stations (Feng et al. 2024; Goodison et al. 1998; Kochendorfer et al. 2022). A recent study by Feng et al. (2024) has shown that CaPA’s solid precipitation analyses could be improved by relaxing its quality control criteria to allow for the assimilation of a greater number of surface observations and by adjusting these observations using a universal transfer function (Kochendorfer et al. 2018; Smith et al. 2020).

Weather radar observations are currently assimilated into CaPA, but only for liquid precipitation. The Canadian radars have been recently equipped with dual-polarization technology, which allows them to identify the type of hydrometeor and provide precipitation estimates for all phases of precipitation (Kumjian et al. 2022). There are several challenges with ground-based radars which might limit their usefulness and impact for wintertime precipitation analysis. The relationship between measured reflectivity and snowfall at the surface remains uncertain due to the intricate microphysics involved in these processes (Matrosov 1997). Additionally, radar reflectivity is valid at a certain height above the surface, even for low pointing angles, necessitating corrections to the vertical reflectivity profile, which relies on certain assumptions (Koistinen and Pohjola 2014). The correction of attenuation and ground clutter is also necessary, as these block the radar beam and cause additional uncertainties (Vivekanandan et al. 1999). Improvements to radar QPEs are possible with radar–gauge adjustment methods, which have been proven useful to mitigate these errors (Fortin et al. 2015; Gjertsen et al. 2003; Goudenhoofdt and Delobbe 2009; Hubbert et al. 2009).

Advances in polarimetric measurements and bias correction of radar QPEs have been shown to improve the quality of snowfall estimates (Ryzhkov et al. 2022; Fassnacht et al. 1999) and are the motivation for testing the inclusion of available radar-based snowfall products in CaPA. The main objective of this study is to determine whether the assimilation of QPEs from weather radars in CaPA could lead to improved solid precipitation estimates over parts of western and central Canada and evaluate the sensitivity of precipitation analyses to the number of surface observations used to correct radar QPE bias.

The paper is organized as follows. Section 2 provides information on the precipitation networks assimilated in CaPA and presents a brief overview of the CaPA system as well as the procedure used to assimilate wintertime radar data. This section also describes the different experiments performed and the evaluation procedures. Section 3 contains the evaluation results, which are further discussed in section 4. Section 5 provides conclusions and future work recommendations.

2. Data and methodology

The analysis domain focuses on the Canadian Prairies and covers the provinces of Alberta, Saskatchewan, and Manitoba at 2.5-km grid spacing (Fig. 1). This domain is smaller than what is used at ECCC for its operations, with a nationwide...
domain for its high-resolution (2.5-km) deterministic and ensemble precipitation analyses (Fortin et al. 2018; section 2, and Khedhaouiria et al. 2020; see Fig. 1). This smaller study domain allows to focus on the center of Canada, which covers a homogeneous domain where snowfall is present.

a. The CaPA system

The precipitation analyses from CaPA are obtained by combining a background $B$ (or first guess) field with observations $O$. This combination is performed using a two-dimensional optimal statistical interpolation technique, specifically a simple kriging of innovations (also known as simple residual kriging or optimal interpolation) (Creutin et al. 1988; Evans 2013). This is done based on the difference between a first-guess value, which is the background field built from a short-range forecast issued by a NWP model, and the corresponding observation. The innovation $[\phi(O_k) - \phi(B_k)]$ is weighted by $W_k$, where $k$ refers to the neighboring observation points. The interpolation procedure, performed on cubic root transforms $[\phi(\cdot)]$, is presented in Eq. (1) and is further explained in Fortin et al. (2018), Khedhaouiria et al. (2020), Lespinas et al. (2015), and Mahfouf et al. (2007):

$$\phi(A) = \phi(B) + \sum_{k=1}^{K} W_k [\phi(O_k) - \phi(B_k)],$$

(1)

where $W_k$, the weights, are obtained through minimizing the variance of the analysis error (Daley 1991). The observation errors from surface stations are assumed to be unbiased and independent, which leads to a diagonal covariance error matrix of variance $\sigma^2_O$. For the ground-based radar data, an exponential and isotropic decline in error with increasing distance from each radar pixel's location is adopted, with a specific error variance $\sigma^2_R$ and correlation length $l_R$. A similar assumption is made for the first-guess spatial error structure, with specific error variance $\sigma^2_B$ and correlation length $l_B$. Error parameters related to both the initial forecast and observations are derived from a variographic approach, independently.

In this study, the first guess comes from a 2.5-km high-resolution configuration of the Global Environmental Multiscale (GEM) model, used as ECCC’s High-Resolution Deterministic Precipitation System (HRDPS, ECCC; Milbrandt et al. 2016). The corresponding version of CaPA at 2.5 km combines observations from surface weather stations and ground-based radars (see section 2b) with the first guess to generate 6- and 24-h precipitation accumulations in liquid water equivalent. This study focuses solely on 6-h analyses for reasons of computational cost, but the results obtained in this study should be validated for 24-h analyses. It should be mentioned that Integrated Multi-satellitE Retrievals for the Global Precipitation Measurement (IMERG) mission (Huffman et al. 2020) satellite products are assimilated in the North American 10-km CaPA version but are not used here.

Prior to the assimilation in CaPA, the ground-based radar product is adjusted in a single-site polar quality control optic to mitigate known issues linked with their retrievals. The four-step procedure followed to achieve this is described in detail in Fortin et al. (2015). First, evident artifacts blocking the radars’ signal are removed using ground stations and dynamic masks for ground clutter removal. Second, a validation of the precipitation pixels reported by the radars is performed using a mask based on Geostationary Operational Environmental Satellite (GOES; Goodman et al. 2012) imagery. This procedure removes radar pixels where there are no clouds reported by the satellites. Third, the radar’s bias is mitigated by conducting a statistical adjustment using surface stations. For each radar, a mean-field bias is estimated by considering the ratio between the sum of radar QPEs and the sum of observed precipitation from surface stations. Only valid radar QPEs close to valid observations from surface stations are taken into account in calculating these sums. In addition, the mean-field bias is calculated for different precipitation observation ranges (0.2–1, 1–3, 3–5, 5–10, 10–20, 20–40, and 40–100 mm). The mean-field bias is adjusted in a recurrent manner by incorporating the previous mean-field biases through a smooth average. The relaxation process involves a continuous series of 150 6-h precipitation cases. Finally, the mean-field bias is applied as a multiplicative correction to each radar’s QPE. This can lead to uncertainties since the ground observations are not necessarily the most reliable and already have errors themselves (Fortin et al. 2015). A final plausibility check removes radar pixels where precipitation retrievals show large differences from the background precipitation. In other words, should the mean absolute error of the normalized radar QPE surpass the mean absolute error of the background on average for the radar’s cells, that entire radar’s data are excluded for that specific time step.

The radar data postprocessing is as described in Fortin et al. (2015), except for the spatial aggregation of radar pixels (Fortin et al. 2015; their section 2) which is performed here on the HRDPS 2.5-km grid (see Fig. 1). Only radar estimates within a 120-km range are assimilated in CaPA to limit errors associated with range. Their assimilation is performed in a similar way to the weather station observations’ assimilation, with the main difference being the consideration of horizontal correlations in the radar data, detailed in Fortin et al. (2015; their section 2).

b. Surface precipitation networks

Observations from various ground-based networks are either assimilated in CaPA or used to objectively evaluate its precipitation analyses. These include observations from station and radar networks, as described below.

1) SURFACE STATION NETWORKS

The surface observations assimilated in CaPA are from multiple weather networks described in Lespinas et al. (2015). They are a collection of manual and automatic stations in Canada and the United States, from networks such as ECCC’s weather stations, the Ontario Ministry of Natural Resources and Forestry (OMNRF), the British Columbia Wildfire Service (BC Ministry of Forests), and the American Meteorological Terminal Aviation Routine (METAR) networks. Further information on the station networks assimilated into CaPA
Table 1. Skill scores (based on contingency tables) with their references and equations. The skill scores depend on the four variables: \(a\) is the number of true-positive events, \(b\) is the number of false-positive events, \(c\) is the number of false-negative events, and \(d\) is the number of true-negative events. The variable \(n\) represents the total and is the addition of \(a, b, c,\) and \(d\).

<table>
<thead>
<tr>
<th>Skill score</th>
<th>Reference</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>FBI</td>
<td>Donaldson et al. (1975)</td>
<td>FBI = (\frac{(a + b)}{(a + c)})</td>
</tr>
<tr>
<td>POD</td>
<td>Donaldson et al. (1975)</td>
<td>POD = (\frac{a}{(a + c)})</td>
</tr>
<tr>
<td>FAR</td>
<td>Donaldson et al. (1975)</td>
<td>FAR = (\frac{b}{(a + b)})</td>
</tr>
<tr>
<td>ETS</td>
<td>Mason (2003)</td>
<td>ETS = (\frac{(a + a_r)}{(a + b + c - a_r)}; a_r = \frac{(a + b)(a + c)}{n})</td>
</tr>
</tbody>
</table>

The snow relations to precipitation rate (cm h\(^{-1}\)) are presented in Eq. (6) (Bukovič et al. 2018; Ryzhkov and Zrnić 1998a,b) for the melting layer and Eq. (7) (Sekhon and Srivastava 1970):

\[
S(Z, K_{DP}) = 1.53K_{DP}^{0.61}10^{-0.33Z},
\]

\[
Z = 39952^{2.21}.
\]

It should be noted that Eq. (6) has been developed for wet snow in the melting layer, which tends to give a strong underestimation of snow precipitation intensity.

3) ADJDLRS DATASET

The Adjusted Daily Rainfall and Snowfall (AdjDlyRS; Mekis and Vincent 2011; Wang et al. 2017) dataset is used in this study for the evaluation of CaPA. This dataset includes adjusted daily precipitation amounts for both rain and snow measurements. Snow depth measurements are performed manually using a ruler and are then converted to liquid water equivalents based on a snow-water conversion map (Mekis and Brown 2010). The rain measurements are from rain gauges, which are corrected for several factors (i.e., wind and wetting loss). Further information on the correction of the data is provided in Wang et al. (2017). The daily precipitation values are valid at 1200 UTC.

c. Evaluation procedures

The quality of CaPA’s precipitation analyses is assessed through selected case studies and for two winter seasons using various statistics presented in Table 1. The evaluation presented in this paper is for two winter seasons, i.e., 2020–21 and 2021–22, from 1 December to 31 March each year. This test period is limited to these 2 years because the new dual-polarized radar-based products for snow are only available since 2020.

First, case studies are useful for obtaining a preliminary and qualitative assessment of the impact of assimilating the DPOPE radar-based product on solid precipitation estimates derived from CaPA’s analyses. Maps are created to compare the precipitation product with and without the radars against an independent set of surface observations (AdjDlyRS). The 6-h accumulated precipitation from CaPA is aggregated over 24-h periods, ending at 1200 UTC. The AdjDlyRS stations
with an altitude over 1000 m are not considered in the evaluation. A bilinear spatial interpolation is performed to obtain CaPA’s precipitation amounts at AdjDlyRS station locations.

Second, it is useful to gain an objective assessment of the impact of assimilating solid precipitation from radars in CaPA. An objective evaluation of precipitation analyses is challenging since most of the precipitation observations available are already being assimilated into the system. For this reason, a leave-one-out cross-validation (LOO CV) approach is used (Feng et al. 2024; Fortin et al. 2015; Friesen et al. 2017). The idea is to use the most reliable stations assimilated into CaPA (i.e., manual synoptic stations), remove them one by one from the observation dataset, and estimate the analysis value at their locations (LOO CV analysis value) using the closest observations. The observation and LOO CV analysis values are then used to compute the verification scores presented below.

The verification of the CaPA product is carried out using the metrics defined in Table 1. The frequency bias index (FBI) is a measure of the bias, with a perfect score of 1. In this study, the values of FBI-1 are shown. Positive values of FBI-1 indicate that CaPA overestimates precipitation event frequency for that specific category, and the opposite for negative values of FBI-1. The probability of detection (POD) shows the hit rate, with a range from 0 (no detection) to 1 (perfect detection). This metric is sensitive to missed events and can be improved by overforecasting. The false alarm ratio (FAR) provides information on the false alarms, with a range from 0 (no false alarms) to 1 (all false alarms); this metric can be improved by underforecasting. To increase detection (POD) and decrease false alarms (FAR) simultaneously is difficult to achieve and should be viewed as an indication of a strong improvement. The equitable threat score (ETS) combines information from both POD and FAR to provide a measure of the accuracy of CaPA precipitation estimates. It ranges from $-1/3$ (worse score) to 1 (perfect score) and is also sensitive to the climatological frequency of events. In general, the interpretation of these metrics is key information for the evaluation as they can all have biases and are only useful as indicators (Pérez Hortal and Michelson 2023).

These four metrics are presented in two manners in this study. First, the metrics are categorized based on the 6-h accumulated precipitation thresholds over the analysis domain for the entire winter season to determine analysis quality as a function of precipitation intensity. With this approach, a non-parametric bootstrap method is performed to statistically identify significant differences between metrics’ values of two experiments, using a 95% confidence interval (Lespinas et al. 2015). Second, to assess the quality of CaPA’s analyses over time, the four metrics are divided into categories of precipitation intensity and displayed in weekly increments throughout the study period.

Finally, partial means (PM) of precipitation are used to inform CaPA’s ability to estimate precipitation accumulation at different intensities based on

$$\bar{x}(\tau) = \langle x | x < \tau \rangle,$$  \hfill (8)

where $x$ is the observed or analyzed (LOO CV) precipitation value and $\tau$ is the precipitation threshold (mm). The observations and analysis values are then compared graphically for multiple thresholds, providing the total precipitation mass as the asymptote for large thresholds.

d. Experimental setup

A series of four experiments were performed for this project, as presented in Table 2. These experiments include the assimilation of radar QPEs for solid precipitation tested in two different CaPA configurations. The first set of experiments (named “Control” and “Radar”) is conducted to evaluate the impact of assimilating radars’ QPE during snowfall in CaPA. The Control configuration is similar to what is currently used operationally at ECCC for its high-resolution deterministic precipitation analysis (Khediaouriia et al. 2020), with the exception of the analysis domain. In Control, CaPA assimilates the QPEs of given weather radars only when the mean surface temperature within the radar domain is above 1.5°C. This means some solid precipitation could be present in the Control experiment, but it remains somewhat negligible for temperatures above 1.5°C. Results from this configuration are compared with the Radar experiment in which radar estimates for both rainfall and snowfall, hence, at all air temperatures. The comparison of these two experiments allows to identify how the assimilation process improves snowfall estimates in CaPA.

Both configurations limit the assimilation of inaccurate precipitation measurements from automatic surface stations. This is because the observations from automatic precipitation gauges suffer from uncertainties associated with undercatch of snowfall in windy conditions (Milewska et al. 2019; Rasmussen et al. 2012). To account for this, CaPA’s operational version only assimilates surface observations from automatic stations when the 2-m wind is less than 0.6 m s$^{-1}$ if the near-surface air temperature is less than 0°C.

The second set of experiments (Control-TF and Radar-TF) aims to assess the impact of assimilating radars when a larger number of surface observations are assimilated. In these
experiments, the wind speed threshold is increased to $3 \text{ m s}^{-1}$ and the universal transfer function (UTF; Kochendorfer et al. 2017) is applied when the measured 2-m temperature is less than $-2^\circ \text{C}$. The larger wind speed threshold allows for an increase of more than 50% the number of observations as-similated in CaPA (Feng et al. 2024). Including more surface observations in the analysis, however, leads to the assimilation of more underestimated precipitation amounts because of the precipitation gauge wind undercatch. The UTF is then used to adjust solid precipitation amounts and is written as follows:

$$CE = \exp[-a(U)\{1 - \tan^{-1}[b(T_{\text{air}})] + c\}], \quad (12)$$

where $a$, $b$, and $c$ are coefficients defined in Kochendorfer et al. (2017), $U$ is the wind speed at 2 m above the ground, and $T_{\text{air}}$ is the air temperature at that same level.

The Control and the Control-TF experiments closely resemble those conducted by Feng et al. (2024), although these were performed over a different period and did not assimilate the radar-based product. Given the significant increase in the number of surface stations assimilated in CaPA in Control-TF relative to Control and the impact this modification has on CaPA’s precipitation analyses when increasing the wind speed threshold and with bias correction of the weather radar’s QPE, these changes were also tested in this study. However, surface observations are also used
to calibrate radar QPEs. Therefore, the Control-TF experiment provides more observations to correct radar QPE bias.

3. Results

a. Case studies

The CaPRA precipitation estimates are compared against the AdjDlyRS observations (Figs. 2 and 3) for two case studies. The first event on 27 March 2021 (Fig. 2) exhibits moderate precipitation originating from the western Canada, while the second event that occurred on 18 January 2021 (Fig. 3) displays higher precipitation amounts over the radar network coming from the northwest.

The radars have a small overall impact on the precipitation accumulations for both events. As expected, their impact is local and restricted to the horizontal range of each individual radar. For instance, the radar experiments contribute to a broader distribution of precipitation in the analyses. Wider regions within the domain exhibit precipitation, which can be observed for both cases. Nevertheless, the regions already displaying precipitation in the control experiments do not show any increase. For the moderate first event, the local changes in accumulations are minimal. However, for the more intense second event, the Radar experiment displays a decrease in accumulated precipitation in regions with highest accumulations (>2 mm), primarily under CASRA (Radisson, Saskatchewan) and CASBE (Bethune, Saskatchewan). A similar pattern is observed when comparing the Control-TF to the Radar-TF, although with generally higher accumulated water shown on the map.

The spatial distribution of precipitation is improved using the radars. The radar experiments (Control-TF and Radar-TF)
more accurately represent areas where the AdjDlyRS stations recorded precipitation compared to the control experiments. For example, under the CASFW (Foxwarren, Manitoba) radar (Figs. 2b, d), precipitation accumulations are closer to the observations, whereas no precipitation is seen for the experiments without the radars (Figs. 2a,c), which led to a precipitation underestimation. A similar pattern is produced at CASRA and in the northern region of CASFW for the second event (Fig. 3).

In contrast, the radar experiments somewhat produce a precipitation overestimation in areas where no precipitation was measured by AdjDlyRS. This is notable under CASBE for the first event, where the control experiment and AdjDlyRS stations indicate no accumulations, but the radar experiments introduce precipitation, resulting in overestimation for small accumulations. A noteworthy observation is that regions where CaPA accurately produced precipitation in the control experiments maintained this accuracy in the radar experiments, such as under CASDR (Dryden, Ontario) for the first event and CASRA for the second event. The radars thus mostly modified precipitation accumulations in areas with already low accuracy.

In summary, the radar experiments indicate an improvement of the localization of precipitation in areas with precipitation (over 0.2 mm). The representation of areas without precipitation, however, seems to pose challenges for the radar experiments, potentially resulting in overestimation based on these two events. This initial analysis of the results provides an insight into the behavior of CaPA’s analyses when integrating radar data and shows promising outcomes. A comprehensive objective comparison of all experiments is however necessary to draw stronger conclusions.

b. Objective evaluation

This section presents results from the objective evaluation of the experiments mentioned in section 3. They are presented with a comparison of two sets of experiments (one with the radars and one without) to analyze the impact of the radars. The following results are discussed below: Control versus Radar, Control-TF versus Radar-TF, and Control versus Radar-TF (section 3).
The comparison of the two experiments (Control and Radar) in the objective evaluation is shown in Figs. 4–6. The overall impact and the time series are shown.

Figure 4 shows how the different metrics are impacted by the inclusion of radars in CaPA, Control versus Radar, for both winters combined. A statistically significant increase in POD was observed for thresholds of 0.2 and 0.5 mm. A largely lowered FAR over all thresholds is also obtained, and when combined with the higher POD, this result is substantial. Finally, the ETS of the radar experiment is larger than the control experiment, indicating a better agreement between CaPA’s analyses and surface observations when the DPQPE radar-based product is assimilated.

The FBI scores show that the assimilation of the DPQPE radar product leads to an overall decrease in the snowfall frequency in the analyses for all precipitation thresholds. This decrease is statistically significant for thresholds of 0.5, 1.0, and 2.0 mm. This translates to an improvement of the bias for small thresholds (reduced overestimation) and a deterioration of the bias at large thresholds (heightened underestimation).

The partial means (Fig. 5), which represent precipitation biases as a function of intensity, indicate closer to observations total precipitation mass with the radar experiment compared to the control experiment, where the radars’ impact seems to be positively lowering the total precipitation, which is consistent with the FBI-1 results. This improvement in partial mean could be partly attributed to the underestimation of the larger-intensity events with the overestimation of the smaller-intensity events. These biases balance out and improve the total precipitation amount compared to observations. These remarks of lowered precipitation at high intensity align with the results from the case studies (section 3). It is noted that the intensity of precipitation is generally lower with radars, which can render a lower frequency bias and lower precipitation amount.

A general look at the evolution of the metrics through the winters (Figs. 6 and 7) gives a perspective on the variation of CaPA’s performance and the impact of the radars. For instance, the metrics indicate deterioration (lower POD, higher FAR, higher frequency bias, and lowered ETS) for the colder winter months (mid-January to mid-March), especially for the higher precipitation thresholds in both years, but notably for the first winter. This reduction is found for both Control and Radar. The differences in scores between the radar and control experiments are quite constant throughout the two winter seasons for all metrics, indicating that the impact of assimilating radar products is relatively independent of the time of year.

The differences between the two experiments are more sensitive to the precipitation thresholds. To better compare these differences, the percentages of points where the Radar score is closer to the perfect score (as per the definition in section 2) than Control are shown in Figs. 6 and 7. The best improvement is shown for the 0.5-mm precipitation threshold, with 81.75% of improvement and 6.75% of neutrality on average for all metrics, and the lowest one for the 2-mm precipitation threshold, with 57.25% of improvement and 27.5% of neutrality on average. This indicates that the assimilation of radars in CaPA faces more challenges for large thresholds, but that their contribution remains positive.

It should be noted that the sample size decreases as the threshold increases, as higher-intensity events have a lower occurrence. This could explain the noisier data for the higher thresholds (panels of 1 and 2 mm). It also explains the gaps in these data, where no events occurred.

Finally, the average of each metric for the two winter seasons is presented in Figs. 6 and 7. Results indicate that the differences between Control and Radar are relatively small for the lowest 6-hourly precipitation threshold (panels of 0.2 mm), in contrast to greater differences for more intense precipitation events (panels of 1 and 2 mm). This indicates that assimilating radar QPEs leads to more accurate analyses for the precipitation envelope, or the areas with precipitation, since 0.2 mm is considered here as the limit threshold for precipitation detection. This is consistent with what was observed in the case studies (section 3), where the detection of actual precipitation seemed more positively impacted by the experiments assimilating radars.

2) IMPACT OF RADAR DATA IN A CONFIGURATION WITH MORE SURFACE OBSERVATIONS

Surface observations assimilated in CaPA are also used to calibrate radar data. It is therefore possible that changes to these surface observations could have an impact on the assimilation of radars in CaPA. In this section, an additional configuration to the assimilation of radars is tested, in which around 200 more surface-adjusted (with UTF) observations are assimilated each winter. The temporal evaluation of metrics for this configuration shows similar results to what was
already observed in section 3; so it is not shown and only evaluation as a function of precipitation threshold is discussed below.

It should be first pointed out that the Control-TF experiment displays improved scores compared to the Control experiment. The Control-TF in Fig. 8 has higher POD, generally lower FAR, and higher ETS, and is closer to zero FBI-1 than in Fig. 4. This also aligns with the conclusions by Feng et al. (2024) with similar experiments, in which more surface-adjusted gauge observations contributed to improve quality of precipitation analyses produced by CaPA.

A comparison between Control-TF and Radar-TF (Fig. 8) reveals a predominantly positive influence when the radar QPEs are assimilated, even when considering an enhanced CaPA configuration (Control-TF). The evaluation metrics exhibit similar trends to those presented in section 3, but with lower impact. The main difference with the results previously described (Fig. 4) is for FBI-1. With Radar-TF, the bias is increased (worsened) for weak thresholds, while it is still deteriorated for larger thresholds. The partial mean (Fig. 9) also shows a degradation of Radar-TF compared with Control-TF, but the difference is relatively small. In general, however, Fig. 8 still shows substantial improvements for POD (for small threshold), FAR (for larger thresholds), and ETS (for all thresholds) achieved with Radar-TF and indicates a positive impact of radars in an enhanced CaPA configuration.

3) IMPACT OF ALL MODIFICATIONS (SURFACE OBSERVATIONS, UTF, AND RADARS)

To assess the total impact of the various configurations tested in this study, Fig. 10 shows Control compared to Radar-TF. A substantially higher POD, lower FAR, and higher ETS over all thresholds of precipitation are observed.
in Radar-TF. The FBI-1 displays a lowered bias with Radar-TF in comparison with Control for thresholds under 1 mm and an increased negative bias for thresholds larger than 1 mm. Globally, a substantial improvement in CaPA's performance is observed when more surface observations and radar QPEs are assimilated.

Last, Fig. 11 shows the partial means of these configurations, and the precipitation intensity distribution of the Radar-TF experiment is closer to that observed than that of the control experiment. With the consideration of all elements, a large improvement in CaPA's ability to reproduce precipitation estimates from manual synoptic stations can be noted with Radar-TF in comparison with Control.

4. Discussion

The results obtained in this study reveal that the assimilation of radar QPEs has a positive impact on the quality of precipitation analyses produced by CaPA over the subdomain evaluated, except for the slight deterioration in frequency bias observed for moderate to heavy precipitation events. This demonstrates that, despite the difficulties of accurately estimating solid precipitation, weather radars are a valuable additional source of information for CaPA during the winter season.

The deterioration in frequency bias caused by the assimilation of radar QPEs represents an opportunity to improve the handling of the radar data in CaPA, where a slight tendency for overestimation is present at small precipitation thresholds (<1 mm), and a stronger tendency for underestimation at larger precipitation thresholds is noted. Representativeness errors linked to the relatively low station density (Gervais et al. 2014) could also contribute to explaining these errors.

a. Bias evaluation

The mean-field debiasing procedure of the radars in CaPA is further examined to gain an understanding of the increase in frequency bias when radar QPEs are assimilated. First, the possibility of the bias being impacted by the range from the radars is investigated, and second, the bias as a function of the intensity of precipitation is explored.

Studies have shown an attenuation correction of radar QPEs assists with the mitigation of underestimation (Borga et al. 2002; Handoyo et al. 2021; Hiley et al. 2011). To
examine the effect of range on the bias of the radars and evaluate the potential benefit of implementing a range-based correction of the radar QPEs, a study was conducted on the categorized bias as a function of the distance from radar sites. The random behavior of the bias suggests that the distance from the radars does not seem to impact the bias in a manner that could be easily corrected (not shown). It is possible that the radars’ conservative range used in CaPA, 120 km, might not exhibit the range effect of bias, where it is masked by the already existing intensity bias observed previously and noted in section 2. In CaPA’s configurations used in this study, a correction of radar QPEs based on the range would most probably not affect CaPA’s analyses. Extending the range of the radars accepted in CaPA (e.g., to 240 km) to then perform a bias evaluation in terms of range could lead to more significant results.

Another approach to bias correction of radar QPEs is as a function of precipitation intensity. It was shown in section 3 that radars’ impact on CaPA’s bias depends on precipitation intensity. This motivated the investigation of the bias binned in terms of intensity of precipitation (6-h accumulations) for the two radar experiments. Comparing the two experiments for this evaluation only allows us to assess the impact of the number of observations assimilated on the bias estimates, since the radar data remain the same.

The radar biases (radar QPEs/observations) as a function of precipitation intensity are shown in Fig. 12. It is found that the bias ~ 1 is well above 0 for small precipitation intensities (0.2–1 mm), demonstrating that the radar-based product over-estimates the number of precipitation events with quantities

![Fig. 8. As in Fig. 4, but for the comparison between Control-TF and Radar-TF.](image)

![Fig. 9. As in Fig. 5, but for the comparison between Control-TF and Radar-TF.](image)
below 1 mm. Conversely, radar QPEs show a negative bias (bias $< 0$) for precipitation above 1 mm, which tends to increase in magnitude with increasing precipitation quantities. This diagnostic improves understanding of the impact of assimilating radar QPEs on the frequency bias in section 3, where we observed that the addition of radar data contributed to worsening the underestimation of the frequency of moderate to large precipitation events (Figs. 4, 8, and 10). In contrast, this might imply that the addition of radar QPEs removed precipitation in wrong places, which would be supported by the partial means. The underestimation of precipitation amounts in the CaPA analyses for precipitation above 2 mm is also visible in Figs. 5 and 11, where the positive differences between Radar or Radar-TF and Control partial mean diminish with increasing precipitation thresholds. This result is also consistent with the case studies (section 3).

This suggests that a more sophisticated approach to correct the radar QPE biases could lead to further improvement in CaPA’s wintertime analyses. The mean-field bias is lower than the bias associated with precipitation amounts between 0.2 and 1 mm and higher for precipitation amounts greater than 1 mm (Fig. 12). An unbiasing technique more adequate to this observed behavior, such as intensity-dependent unbiasing of radar data, could provide more accurate precipitation estimates from radars in winter and potentially from CaPA.

b. Evaluation limitations

The evaluation presented in this study covered only two winter seasons. Several winter seasons would have been
needed to better assess the impact of assimilating radar QPEs on CaPA’s precipitation analyses. The two winters are, however, different from a meteorological point of view, winter 2020–21 was drier and warmer, and winter 2021–22 was wetter and colder with a higher proportion of solid precipitation events than the preceding winter. Although different, the impact of assimilating radar data is consistent throughout both seasons (section 3). This indicates that similar impacts can be expected for other winters. It also highlights the potential for expanding the usage of radar for solid precipitation in CaPA.

Another limitation concerns the way the LOOCV evaluation is performed. The stations for evaluation are the same as those used for radar unbiasing. However, as concluded in section 3, the correction applied using additional surface stations (Radar-TF) on the radars has a relatively minor effect on CaPA’s analyses. This suggests that the utilization of stations to correct radar biases has minimal impact on the evaluation procedure.

This study was only performed on a fraction of CaPA’s operational domain. Nevertheless, it is a good representation of the snow-covered area of Canada. The CaPA configurations tested in this study should be further evaluated for the whole domain, especially for the mountains of western and eastern Canada. As there are no radars in northern Canada, the impact should be minimal in this remote region.

Finally, winter accumulations from CaPA (Radar-TF experiment) are compared with those from ERA5-Land (Muñoz-Sabater et al. 2021) to gain insights into the differences between the two products. Overall, spatial features from CaPA’s seasonal precipitation are similar to those from ERA5-Land, with larger amounts in the central portion of Alberta, Saskatchewan, and Manitoba and lesser amounts on the lee side of the Rocky Mountains. Both products generally have larger accumulations for the 2021–22 winter season. Precipitation from CaPA, however, displays more spatial variability. Also, CaPA appears to analyze less precipitation than ERA5-Land, as evidenced, for example, by the smaller amounts in central and southern Alberta. It is important to emphasize that a rigorous and extensive comparison of the two products is beyond the scope of this article. Figure 13 is shown only to offer a broader understanding of the similarities and distinctions of CaPA’s precipitation analyses when compared to another established product such as ERA5-Land.

5. Conclusions

In conclusion, the addition of radar QPEs for solid precipitation contributes to significantly improve the quality of CaPA’s analyses for minor precipitation events (<1 mm h⁻¹), with little sensitivity to the number of surface stations assimilated or to their correction by the UTF. The assimilation of radar QPEs for solid precipitation has more mixed effects on the quality of CaPA’s analyses for larger precipitation events (>1 mm h⁻¹), for which radars contribute to their underestimation, again with a very low sensitivity to the number of surface stations and the use of the transfer function. Overall, the significant increase in ETS and POD and the simultaneous decrease in FAR suggest that solid radar QPEs are a valuable complement to surface observations for solid precipitation estimates produced by CaPA.

Additional correction of the radar QPEs in winter based on ground observations could further improve the positive contribution. The current mean-field bias correction performed in CaPA remains widely supported and is one of the most used snowfall bias-adjustment techniques in radar QPEs (Chumchean et al. 2006; Holleman 2007; Wilson 1970). Studies have assessed other promising techniques for using gauge information to correct radar fields for rainfall, with Kalman filtering or copula-based approaches, for instance (Anagnostou et al. 1998; Chumchean et al. 2006; Goudenhoofdt and Delobbe 2009; Vogl et al. 2012). Correction of radar QPEs for mixed precipitation could also be beneficial, instead of only considering rain and snow. Provided that appropriate corrections to radar QPEs are in place, the maximum radar range tolerated for assimilation in CaPA could be increased from 120 km to their operational value of 240 km, thereby substantially increasing the amount of radar data assimilated in CaPA.

Additionally, further studies on other methods to improve CaPA’s winter analyses are being undertaken. Tests aimed at correcting observations from automatic stations with various transfer functions and the assimilation of satellite data to the system, from the Integrated Multi-satellite Retrievals for Global Precipitation Measurement (GPM) (IMERG; Asong et al. 2017; Huffman et al. 2020; Pradhan et al. 2022), will provide a better understanding on the factors that impact CaPA’s...
performance in winter and possibly lead to better solid precipitation estimates.

**Acknowledgments.** The research was made possible through the support of Manitoba Hydro (MH) and Mitacs funding. This research was enabled in part by support provided by Calcul Québec (www.calculquebec.ca) and the Digital Research Alliance of Canada (alliancecan.ca). JMT would also like to acknowledge the Canada Research Chairs Program and the Natural Sciences and Engineering Research Council of Canada. Lastly, we thank all our collaborators from MH (especially Kevin Sagan, Scott Herbert, and Shane Wruth) and ECCC (Eva Mekis, Guy Roy, and Sudesh Boodoo) for their constructive engagement along the project.

**Data availability statement.** All the experiment files of Control, Control-TF, Radar, and Radar-TF generated from CaPA for this study and used for evaluation can be made available upon request.

**REFERENCES**


Evans, A., 2013: Investigation of enhancements to two fundamental components of the statistical interpolation method used by the Canadian Precipitation Analysis (CaPA) M.S. thesis, Dept. of Civil Engineering, University of Manitoba, 307 pp.


Koistinen, J., and H. Pohjola, 2014: Estimation of ground-level reflectivity factor in operational weather radar networks using


