Quantifying the Impact of Precipitation-Type Algorithm Selection on the Representation of Freezing Rain in an Ensemble of Regional Climate Model Simulations

CHRISTOPHER D. MCCRAY,[a,b] JULIE M. THÉRIAULT,[b] DOMINIQUE PAQUIN,[a] AND ÉMILIE BRESSON[a]

[a] *Ouranos, Montréal, Québec, Canada*
[b] *Department of Earth and Atmospheric Sciences, Université du Québec à Montréal, Montréal, Québec, Canada*

(Manuscript received 27 September 2021, in final form 28 March 2022)

ABSTRACT: Given their potentially severe impacts, understanding how freezing rain events may change as the climate changes is of great importance to stakeholders including electrical utility companies and local governments. Identification of freezing rain in climate models requires the use of precipitation-type algorithms, and differences between algorithms may lead to differences in the types of precipitation identified for a given thermodynamic profile. We explore the uncertainty associated with algorithm selection by applying four algorithms (Cantin and Bachand, Baldwin, Ramer, and Bourgouin) offline to an ensemble of simulations of the fifth-generation Canadian Regional Climate Model (CRCM5) at 0.22° grid spacing. First, we examine results for the CRCM5 driven by ERA-Interim reanalysis to analyze how well the algorithms reproduce the recent climatology of freezing rain and how results vary depending on algorithm parameters and the characteristics of available model output. We find that while the Ramer and Baldwin algorithms tend to be better correlated with observations than Cantin and Bachand or Bourgouin, their results are highly sensitive to algorithm parameters and to the number of pressure levels used. We also apply the algorithms to four CRCM5 simulations driven by different global climate models (GCMs) and find that the uncertainty associated with algorithm selection is generally similar to or greater than that associated with choice of driving GCM for the recent past climate. Our results provide guidance for future studies on freezing rain in climate simulations and demonstrate the importance of accounting for uncertainty between algorithms when identifying precipitation type from climate model output.

SIGNIFICANCE STATEMENT: Freezing rain events and ice storms can have major consequences, including power outages and dangerous road conditions. It is therefore important to understand how climate change might affect the frequency and severity of these events. One source of uncertainty in climate studies of these events is related to the choice of algorithm used to detect freezing rain in model output. We compare the frequency of freezing rain identified using four different algorithms and find sometimes large differences depending on the algorithm chosen over some regions. Our findings highlight the importance of taking this source of uncertainty into account and will provide researchers with guidance as to which algorithms are best suited for climate studies of freezing rain.

KEYWORDS: Climate variability; Freezing precipitation; Mixed precipitation; Downscaling; Climate models

1. Introduction

Freezing rain is a hazardous winter weather phenomenon that has resulted in billions of dollars in damage in Canada and the United States (e.g., Changnon 2003). Electrical utility companies, for example, are especially impacted by these events as the accretion of ice on trees and power lines can damage electrical infrastructure and result in prolonged outages. Given these impacts, understanding how the frequency and severity of ice storms may change as the climate changes is of great importance to stakeholders. Though several studies have examined these events in climate model output, a great deal of uncertainty remains in these projections. One of the key uncertainties that has yet to be fully explored is that associated with the method used to identify freezing rain in model output, namely the selection of precipitation-type algorithms that use differing criteria to detect freezing rain.

Two mechanisms can lead to freezing rain formation. The classical mechanism, known as the melting process (e.g., Brooks 1920; Meisinger 1920), involves hydrometeors that begin as snow aloft at altitudes where temperatures are sufficiently cold for ice nucleation. These snowflakes then fall into an above-freezing warm layer that is sufficiently warm/deep that they completely melt. The raindrops then fall into a below-freezing surface-based cold layer and freeze on contact at the surface. The second mechanism, known as the supercooled warm rain process (Huffman and Norman 1988; Rauber et al. 2000), does not require the presence of a warm layer aloft. Precipitation forms as supercooled liquid drops in a saturated layer too warm for ice nuclei to be active. This process often occurs with shallow saturated layers near the surface and thus often results in freezing drizzle, though freezing rain is also possible (Rauber et al. 2000).

In North America, the conditions necessary for freezing rain formation occur most frequently in the eastern United

Denotes content that is immediately available upon publication as open access.

Corresponding author: Christopher D. McCray, mccray.christopher@ouranos.ca

DOI: 10.1175/JAMC-D-21-0202.1

States and Canada (e.g., Cortinas et al. 2004; McCray et al. 2019), and these conditions are heavily influenced by mesoscale meteorological conditions associated with terrain features (e.g., Bernstein 2000; Razy et al. 2011). Maximum freezing rain frequency occurs in eastern Newfoundland (median of ≥50 h yr^{-1}) and in the Saint Lawrence River valley of Québec, Canada (40–50 h yr^{-1}). In the latter region, cold air channeled down the valley sustains the surface-based cold layer and supports prolonged freezing rain events (Carrera et al. 2009; Razy et al. 2011; Ressler et al. 2012). An extreme example is the catastrophic January 1998 ice storm that remains one of Canada's costliest natural disasters (Gyakum and Roebber 2001; Roebber and Gyakum 2003; Henson and Stewart 2007). In western North America, freezing rain also regularly occurs in portions of Alaska and the Columbia Basin region of the Pacific Northwest (median of 5–10 h yr^{-1}) (Cortinas et al. 2004; McCray et al. 2019).

In part because of the importance of local terrain features on their occurrence, simulation and prediction of freezing rain events remains a challenge (e.g., Ralph et al. 2005). Some modern high-resolution numerical weather prediction (NWP) models can explicitly diagnose precipitation type using hydrometeor mixing ratios from the model microphysics package (e.g., Benjamin et al. 2016). Because of computational constraints, precipitation type is generally not explicitly diagnosed in climate models, and freezing rain must be identified using precipitation-type algorithms. These algorithms were initially developed for NWP models and use variables including geopotential height, temperature, and humidity to determine the expected phase of precipitation at the surface (e.g., Cantin and Bachand 1993; Baldwin and Contorno 1993; Ramer 1993; Czys et al. 1996; Bourgouin 2000). Differences among the algorithms, for example thresholds used to distinguish precipitation types, can lead to differences in the phase of precipitation identified for a given thermodynamic profile. Reeves et al. (2014) verified five algorithms against observations and found relatively poor skill for the identification of freezing rain compared with rain or snow, with algorithms having difficulty distinguishing freezing rain from ice pellets.

To accurately reproduce the conditions necessary for freezing rain, models must be of sufficiently high resolution to resolve key topographic features. Lambert and Hansen (2011) examined freezing rain in a global climate model (GCM), CGCM3, with 2.8° grid spacing. The coarse resolution of this simulation resulted in smoothed topography and produced a climatology that lacked the local maxima in freezing rain frequency over the northwestern United States and in the Saint Lawrence River valley. Cholette et al. (2015) compared five simulations of differing horizontal resolution over the Saint Lawrence River valley and found that at 0.81°, wind channeling and the vertical temperature profile responsible for freezing rain were poorly reproduced. The highest resolution 0.03° and 0.01° simulations best represented the terrain and associated meteorological impacts. Recent studies have therefore employed higher-resolution regional climate models (RCMs) that better represent terrain features, driving the RCM with reanalysis datasets to examine the past climate or GCMs to explore future changes.

Several studies have examined the ability of the fifth-generation Canadian Regional Climate Model (CRCM5; Martynov et al. 2013; Šeparović et al. 2013) driven by ERA-Interim (Dee et al. 2011) to reproduce the observed freezing rain climatology. Bresson et al. (2017) found that at high resolution (0.11°), the CRCM5 accurately reproduced the spatial pattern of the freezing rain climatology over eastern Canada but overestimated freezing rain frequency. St-Pierre et al. (2019) found that CRCM5 simulations at 0.11° and 0.22° overestimated freezing rain frequency but better reproduced individual events and the local maxima in the freezing rain climatology than a 0.44° simulation. Other studies have driven RCMs with GCMs to analyze future projections of freezing rain (Matte et al. 2019; Kämäräinen et al. 2018; Jeong et al. 2018, 2019).

A key limitation of existing studies on freezing rain in climate simulations is that they have generally chosen a single precipitation-type algorithm, and uncertainty associated with algorithm selection has not been fully explored. For example, Lambert and Hansen (2011) applied the Ramer (1993) algorithm while Bresson et al. (2017), Jeong et al. (2018), St-Pierre et al. (2019), and Jeong et al. (2019) used the Bourgouin (2000) method. Some studies have applied algorithms online as the model runs at each time step, taking advantage of all vertical model levels (e.g., Bresson et al. 2017; St-Pierre et al. 2019). In most cases, freezing rain is not among archived model variables and must therefore be identified offline via postprocessing of available model output typically archived on fewer pressure levels and/or at a reduced frequency (e.g., Lambert and Hansen 2011; Matte et al. 2019; Jeong et al. 2019).

Comparisons between algorithms have primarily been undertaken with reanalysis datasets. Mullens and McPherson (2017) assessed the ability of the North American Regional Reanalysis (NARR; Mesinger et al. 2006) to reproduce the climatology of freezing rain over the south-central United States. They compared results obtained from NARR's online freezing rain product [generated using the Baldwin and Contorno (1993) algorithm] and those obtained by applying three algorithms (Cantin and Bachand 1993; Baldwin and Contorno 1993; Bourgouin 2000) offline. Both the offline and online techniques were able to reliably reproduce individual ice storms and the climatology of freezing precipitation, though the three-algorithm ensemble mean produced more frequent freezing precipitation than the online method.

Jeong et al. (2019) identified freezing precipitation changes over North America in a 50-member ensemble of fourth-generation Canadian RCM (CanRCM4; Scinocca et al. 2016) simulations driven by CanESM2 (Arora et al. 2011). Although they only applied the Bourgouin (2000) algorithm offline to model output, they also examined freezing rain frequency and quantity in NARR and found that the online Baldwin calculation identified freezing rain more frequently than the Bourgouin (2000) algorithm applied offline.

Matte et al. (2019) was the first study to present an analysis of climate model output using multiple precipitation-type algorithms. They examined occurrences of mixed precipitation (freezing rain and ice pellets) over southern Québec, Canada, using five algorithms (Cantin and Bachand 1993;

Baldwin and Contorno 1993; Ramer 1993; Czys et al. 1996; Bourgouin 2000) applied offline to 0.11° simulations of CRCM5 and compared algorithms using the hindcast simulation driven by ERA-40 (Uppala et al. 2005) and ERA-Interim (Dee et al. 2011). They found a broad variation of mean annual mixed precipitation frequencies depending on algorithm, with values at Montréal (CYUL) ranging from ~20 h yr^{-1} with Cantin and Bachand (1993) to ~45 h yr^{-1} with Baldwin and Contorno (1993).

Given the differences between algorithms found by Matte et al. (2019) in CRCM5 over southern Québec and by Mullens and McPherson (2017) and Jeong et al. (2019) using NARR, a more complete understanding of the impact of precipitation-type algorithm selection on the representation of freezing rain in climate models is necessary. The objective of this paper is therefore to quantify the uncertainty associated with algorithm selection, in particular relative to other sources of variability. To do this, we identify freezing rain in an ensemble of CRCM5 simulations over North America using four precipitation-type algorithms for the recent past climate. We first compare how well each algorithm reproduces the observed climatology of freezing rain using a reanalysis-driven simulation. We further examine sensitivity of the simulated climatology to choices related to the identification of freezing rain in model output, including choices for algorithm parameters, minimum precipitation thresholds, and the number of vertical levels available. Finally, we compare the variability between different algorithms to 1) variability among simulations produced using different driving GCMs and 2) natural variability, estimated using the CRCM5 driven by multiple members of a single GCM.

2. Data and methods

a. Climate simulations

We examine a suite of CRCM5 simulations at 0.22° grid spacing within the North American component of the Coordinated Regional Climate Downscaling Experiment (NA-CORDEX; Mearns et al. 2017). The CRCM5 at this resolution has been shown capable of reproducing the climatology of freezing rain over northeastern North America (St-Pierre et al. 2019). We first analyze freezing rain in the evaluation simulation driven at its lateral boundaries by the 0.75°, 6-hourly ERA-Interim reanalysis (Dee et al. 2011) for 1979–2016 (herein CRCM5–ERA-Interim). We then examine simulations driven by four different GCMs from 1950 to 2005 for the historical period and 2006–2100 for the future climate: CanESM2 (Arora et al. 2011), MPI-ESM-LR (Giorgetta et al. 2013), GFDL-ESM2M (Dunne et al. 2012), and CNRM-CM5 (Voldoire et al. 2013) (Table 1). All simulations are referred to by the combination of the RCM and driving reanalysis/GCM (e.g., CRCM5–CNRM-CM5). Simulations driven by four additional CanESM2 members with randomly perturbed initial conditions are compared to provide an estimate of natural variability.

All simulations are run with 56 hybrid levels and a 10-min time step. Temperature and specific humidity are archived every three hours at 2 m and on 17–27 pressure levels depending on the simulation (Table 1). Vertical resolution is at least every 25 hPa from 1000 to 850 hPa and generally every 50–100 hPa at lower pressures, with additional levels available for some simulations. We examine the 1980–2009 period, with data for 1980–2005 from the historical simulations using observed greenhouse gas concentrations and 2006–09 taken from the future simulations forced with the RCP 8.5 emissions scenario. Additional details on the CRCM5 and its parameterizations can be found in St-Pierre et al. (2019).

TABLE 1. List of GCMs and reanalysis datasets used to drive the CRCM5 simulations used in this study along with the GCM ensemble member used and the number of pressure levels archived for each simulation.

Driving reanalysis/GCM	Ensemble member	*n* archived pressure levels
ERA-Interim	—	27
CNRM-CM5	r1i1p1	27
MPI-ESM-LR	r1i1p1	17
GFDL-ESM2M	r1i1p1	27
CanESM2	r1i1p1	22
CanESM2	r2i1p1	22
CanESM2	r3i1p1	22
CanESM2	r4i1p1	22
CanESM2	r5i1p1	22

b. Precipitation-type algorithms

We diagnose freezing rain in model output using four algorithms of varying complexity. Identifying precipitation type via postprocessing of model thermodynamic fields poses several challenges. A majority of freezing rain events last fewer than three hours (Cortinas et al. 2004; McCray et al. 2019), and use of three-hourly model output is likely to lead to an underestimation of brief events. Additionally, three-hourly amounts that may include precipitation of multiple phases can only be assigned the precipitation type diagnosed from the thermodynamic profile at the archived time step. Despite these limitations, past studies using climate model or reanalysis datasets with similar temporal and vertical resolution have had relative success at reproducing the North American climatology of freezing rain using some of these algorithms (e.g., Mullens and McPherson 2017; Matte et al. 2019; Jeong et al. 2019). The algorithms used here are chosen either for their simplicity (Cantin and Bachand 1993) or because they are used operationally in U.S. and Canadian NWP models (Ramer 1993; Baldwin and Contorno 1993; Bourgouin 2000).

The four algorithms used to identify freezing rain are applied at grid points where the simulated precipitation rate is at least 1 mm day^{-1} [0.125 mm (3 h) $^{-1}$]. This threshold is used to remove very light precipitation amounts that tend to be overestimated in RCM output and was employed in several past studies using the CRCM5 (Bresson et al. 2017; Matte et al. 2019; St-Pierre et al. 2019). Sensitivity of results to this threshold is explored in section 4b. All presented calculations are restricted to grid points with land fraction ≥ 85%. Each freezing rain occurrence is multiplied by three and results are presented as hours of freezing rain for comparison with

hourly observations, as described in the following section. The algorithms used here include the following:

1) CANTIN AND BACHAND (1993)

The simplest of the four methods we apply is Cantin and Bachand (1993), also known as the partial thickness method, which distinguishes precipitation type based on the thickness of the 1000–850-hPa and 850–700-hPa layers. This technique requires only geopotential height on three pressure levels, allowing it to be applied to simulations lacking detailed temperature and humidity output on multiple vertical levels. Though Cantin and Bachand (1993) provides detailed thickness thresholds based on the qualitative synoptic pattern, we use a simplified method as in Matte et al. (2019). When the 850–700-hPa layer thickness $\geq$ 154 dam, melting is assumed to occur aloft. If the 1000–850-hPa thickness is less than 129 dam, refreezing is expected and ice pellets are diagnosed, whereas values of 129–131 dam are associated with freezing rain. Rain is diagnosed when the 850–700-hPa layer thickness $\geq$ 154 dam and 1000–850-hPa thickness is greater than 131 dam. As in Matte et al. (2019), we also add a requirement that surface temperature $T_s \leq 0$°C for freezing rain.

The Cantin and Bachand method was developed using data over southern Québec, Canada, where 850 and 700 hPa are typically representative of the above-freezing warm layer during freezing rain. Over higher terrain, for example, the Rocky Mountains, the 1000-, 850-, and even 700-hPa levels may be below the surface. Though we present results obtained using Cantin and Bachand at all grid points, the model geopotential height data used is extrapolated in cases where the given level is below ground. Some studies have modified Cantin and Bachand to use levels and thresholds more appropriate for higher terrain locations (e.g., Mullens and McPherson 2017), though this requires changes specific to individual regions that are beyond the scope of our work.

2) BOURGOUIN (2000)

The Bourgouin (2000) method involves a comparison of the positive and negative areas (PA and NA, respectively) surrounding the 0°C isotherm on the vertical dry-bulb temperature profile. These areas are calculated using the thickness and mean temperature of each layer, with thresholds for precipitation types identified using two seasons of observed surface and upper-air data over North America. If the surface-based layer has PA ≤ 5.6 J kg^{-1} or the PA above a surface-based cold layer is less than 2 J kg^{-1}, snow is diagnosed. Otherwise, freezing rain is identified if NA $< 46 + 0.66 \times$ PA, indicative of a sufficiently warm/deep warm layer for melting but an insufficiently cold/deep cold layer for refreezing. If NA $> 66 + 0.66 \times$ PA, ice pellets are diagnosed, and if $46 + 0.66 \times \text{PA} \leq \text{NA} \leq 66 + 0.66 \times$ PA, a mixture of freezing rain and ice pellets is possible.

The Bourgouin algorithm has two key limitations. First, it uses dry-bulb rather than wet-bulb temperature, and therefore does not consider the effects of evaporative cooling of hydrometeors. Second, it has no criteria associated with the vertical moisture profile. Thus, a vertical temperature profile conducive to freezing rain that is completely unsaturated could be attributed to freezing rain despite no precipitation forming. This issue is partly mitigated here as we apply algorithms only at grid points where precipitation was produced in the model. Still, as with Cantin and Bachand, freezing precipitation formed through the supercooled warm rain process with no warm layer aloft cannot be identified using Bourgouin. A revised version of the Bourgouin algorithm addressing these limitations has recently been presented (Birk et al. 2021), although we use the original operational version here.

3) RAMER (1993)

The Ramer (1993) algorithm is a decision-tree approach based on the presence of saturated layers aloft and a calculation of ice fraction I, ranging from 0 (completely liquid) to 1 (completely solid), as precipitation descends toward the surface. First, the precipitation generation level is identified as the first layer below 400 hPa with relative humidity RH greater than a given threshold RH_{crit}, set to 90% in Ramer (1993). We set $\text{RH}_{\text{crit}} = 80\%$ as in the Unified Post Processor code (available online; https://dtcenter.org/community-code/unified-post-processor-upp) used for the National Centers for Environmental Prediction (NCEP) NWP models. We examine sensitivity of results to this choice in section 4. If the wet-bulb temperature $T_w < -6.6$°C throughout the vertical profile, snow is diagnosed at the surface, and if $T_w > 2$°C at the surface, rain is assumed. At the generation level, precipitation begins in solid form and ice fraction I is set to 1 if $T_w < -6.6$°C. Otherwise, precipitation is assumed to begin as liquid and $I = 0$. Freezing rain is diagnosed if precipitation forms as ice, which then melts with $I < 0.04$ and $T_w < 0$°C at the surface. If $0.04 \leq I \leq 0.85$, a mix including freezing rain is identified, which is included in our analysis as an occurrence of freezing rain. Additionally, in cases where $I = 0$ at the precipitation generation level and surface $T_w < 0$°C, freezing rain formed via the supercooled warm rain process is identified. A detailed description of the Ramer method can be found in Cortinas et al. (2002).

The ability of the Ramer algorithm to identify freezing rain formed via the supercooled warm rain process differentiates this method from those of Bourgouin and Cantin and Bachand. However, this capability comes with added sensitivities. Reeves et al. (2014) found that the precipitation type identified is highly sensitive to the ice nucleation threshold $T_w < -6.6$°C, with probability of detection for freezing rain (ice pellets) decreasing (increasing) as this threshold is increased. They also found that Ramer had the best skill scores for freezing rain identification among the five methods they studied, but among the worst scores for ice pellets. Additional uncertainty comes from the threshold for RH_{crit}, analyzed in section 4.

4) BALDWIN AND CONTORNO (1993)

The Baldwin and Contorno (1993, herein Baldwin) algorithm is similar to the Ramer algorithm in that its decision tree contains branches allowing for detection of a precipitation generation layer and identification of freezing precipitation formed through the supercooled warm rain process. For Baldwin, atmospheric layers are considered saturated if the

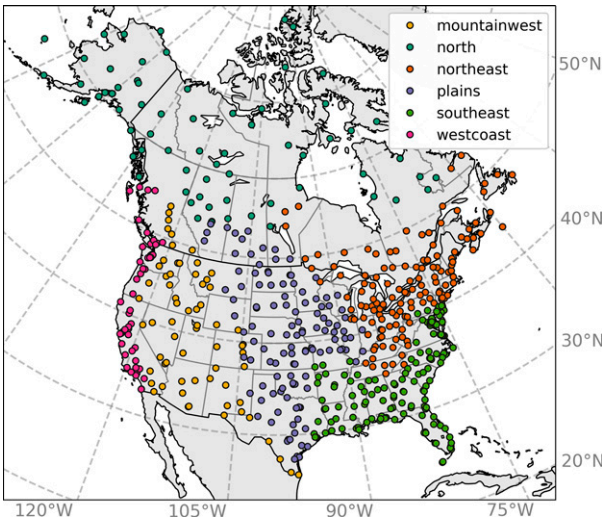

FIG. 1. Map of surface stations used to compare model output and observations. Stations are from the dataset of McCray et al. (2019). Stations are categorized into Bukovsky (2011) regions, which are then grouped into six broader categories indicated by the shading of the points.

dewpoint depression $T_{DD} < 6°C$. In the Unified Post Processor code, a slightly modified version is used whereby saturated layers with $T_{DD} < 2°C$ are first identified. If none are found, a search is done progressively increasing the threshold by 2°C up to the original 6°C threshold. Unless otherwise stated, we use the initial version with a fixed $T_{DD} < 6°C$ threshold and explore sensitivity to the method used in section 4.

The uppermost saturated level is defined as the precipitation generation layer. If $T_w < -4°C$ at this level, precipitation begins as ice, otherwise hydrometeors are initially liquid. Identification of the surface precipitation type involves a comparison of several layer areas (layer depth multiplied by the layer-average T_w). The area between T_w and −4°C from the surface to 500 hPa is used to identify melting associated with warm layers. Freezing rain is diagnosed if precipitation begins as ice, this area is > 3000°C m, surface $T_w \leq 0°C$, and the negative ($T_w < 0°C$) area of the lowest 150-hPa surface-based layer is > −3000°C m. Detailed descriptions of the Baldwin algorithm can be found in Cortinas et al. (2002) and Mullens and McPherson (2017).

The Baldwin method, like Ramer, allows for identification of freezing rain formed via the supercooled warm rain process. This results in similar sensitivities, namely to the $T_w < -4°C$ threshold for ice nucleation and the $T_{DD} < 2°C$ threshold for saturation. Though the supercooled warm rain process tends to produce freezing drizzle, it can also produce freezing rain (Rauber et al. 2000). We therefore follow the original Ramer and Baldwin techniques and past studies that have used them (e.g., Lambert and Hansen 2011; Matte et al. 2019) and identify precipitation as freezing rain regardless of which formation mechanism is responsible.

c. *Observations*

To assess the ability of the simulations to reproduce the climatology of freezing rain, we compare the simulated climatology with the freezing rain observation dataset presented in McCray et al. (2019), which contains hourly observations of freezing rain at 579 surface stations in the United States and Canada from 1979 to 2016. Validation of freezing rain in climate models is challenging as no public long-term dataset of ice accretion exists to compare with simulated precipitation amounts, and frequency of occurrence must be used instead. Surface observations are typically hourly and sometimes subhourly, while model output here is only available every three hours. To address this, we follow Bresson et al. (2017) and group observations into three-hourly periods. For example, if freezing rain is diagnosed at 0100, 0200, and/or 0300 UTC, only one observation of freezing rain is counted, at 0300 UTC. As with model output, each occurrence is multiplied by three and presented as hours of freezing rain.

We categorize stations into the regions presented in Bukovsky (2011) (Fig. 1) to identify regional differences in our results. The 31 Bukovsky regions are grouped into six categories based on presented climatologies of freezing rain (e.g., Cortinas et al. 2004; McCray et al. 2019) and the regions where freezing precipitation occurs under similar synoptic patterns (e.g., Rauber et al. 2001). Stations in Alaska or north of 60°N latitude and outside of the original Bukovsky (2011) regions are assigned the "North" category. The number of stations per category ranges from 47 (West Coast) to 141 (Northeast).

3. Evaluation of the simulated climatology

Figure 2 presents the median annual hours of freezing rain from 1980 to 2009 in the CRCM5–ERA-Interim evaluation simulation for the four algorithms described in section 2b. The

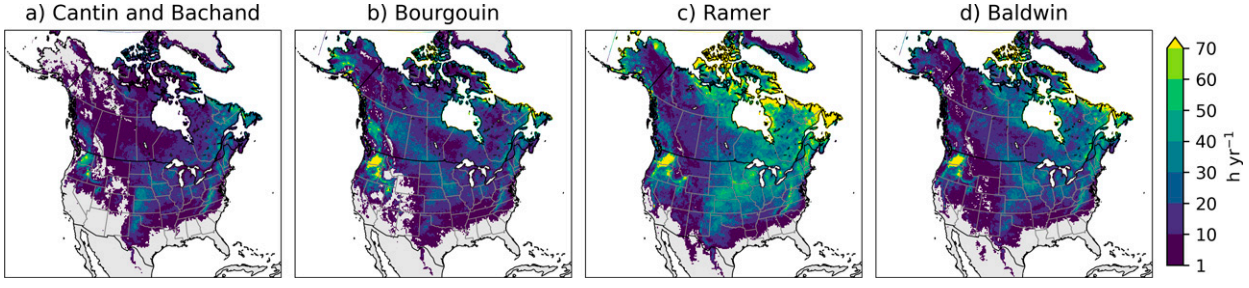

FIG. 2. Median annual hours of freezing rain (1980–2009) in the CRCM5–ERA-Interim simulation calculated using the (a) Cantin and Bachand, (b) Bourgouin, (c) Ramer, and (d) Baldwin algorithms.

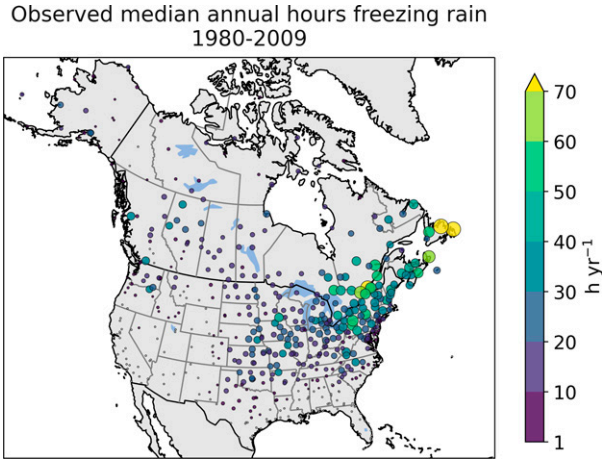

FIG. 3. Observed median annual hours of freezing rain (1980–2009) at the 579 surface stations from McCray et al. (2019), modified for comparison with three-hourly model output as described in section 2c.

spatial pattern is qualitatively similar among algorithms and generally compares well to the observed climatology (Fig. 3), though the magnitude varies substantially among algorithms. All algorithms highlight the observed axis of relatively high freezing rain frequency from Texas northeastward to Newfoundland, and all but Cantin and Bachand reproduce the maximum in the Saint Lawrence River valley. Consistent with Matte et al. (2019), Ramer and Baldwin produce freezing rain most frequently (Figs. 2c,d) followed by Bourgouin (Fig. 2b) and Cantin and Bachand (Fig. 2a).

Pearson correlation coefficients between observed and simulated median annual hours of freezing rain (Fig. 4) range from $r = 0.35$ for Bourgouin (Fig. 4b) to $r = 0.60$ for Ramer (Fig. 4c). Correlation coefficients are reduced for all algorithms due to a cluster of high-biased values at several stations in the North and Mountain West regions (Alaska, Washington, Oregon) with median observed frequencies ≤ 25 h yr^{-1} (Figs. 2, 4). Potential explanations include the CRCM5's structural cold surface temperature bias over these regions and its strong wet bias over Alaska (Šeparović et al. 2013).

Regional differences in the formation mechanisms of freezing rain and the selection of stations used to develop the various algorithms may lead to regional differences in skill between methods. The Ramer algorithm has the best correlation coefficients in four of the six regions but tends to overestimate freezing rain frequency as indicated by its large positive biases (Table 2). Cantin and Bachand has the lowest mean bias in three of the regions, but underestimates freezing rain frequency everywhere but the Mountain West, where some or all of the three pressure levels used are below ground, and therefore data are extrapolated. The best spatial correlations and lowest mean biases are found in the Southeast ($r \geq 0.82$ for all algorithms) followed by the Plains. Skill in these regions is likely aided by the relatively homogeneous terrain compared with the other regions. The poorest correlations and largest biases are found in the north, where $r \leq 0.18$ for all algorithms. This poor skill is possibly related to the very large wet bias in CRCM5 over portions of this region, particularly Alaska and the Canadian Arctic Archipelago (Šeparović et al. 2013).

Rauber et al. (2000) identified the relative frequencies of the melting and supercooled warm rain processes during observations of freezing rain and drizzle over the United States east of the Rocky Mountains. They found freezing precipitation formed via the melting process to be most frequent along and east of the Appalachian Mountains, while the supercooled warm rain process occurred most often over the plains. Freezing precipitation formed via the supercooled warm rain process was most commonly observed as freezing drizzle, while the melting process most commonly produced freezing rain. Over much of Alaska and Canada west of Québec, freezing drizzle is more common than freezing rain (Cortinas et al. 2004). This may partly explain the large positive biases for Ramer and Baldwin over the North and Plains regions (Table 2) as here we are only comparing model output with observations of freezing rain.

Overall, skill varies greatly from region to region, and no single algorithm or configuration performs substantially better than another. While the Ramer and Baldwin algorithms generally

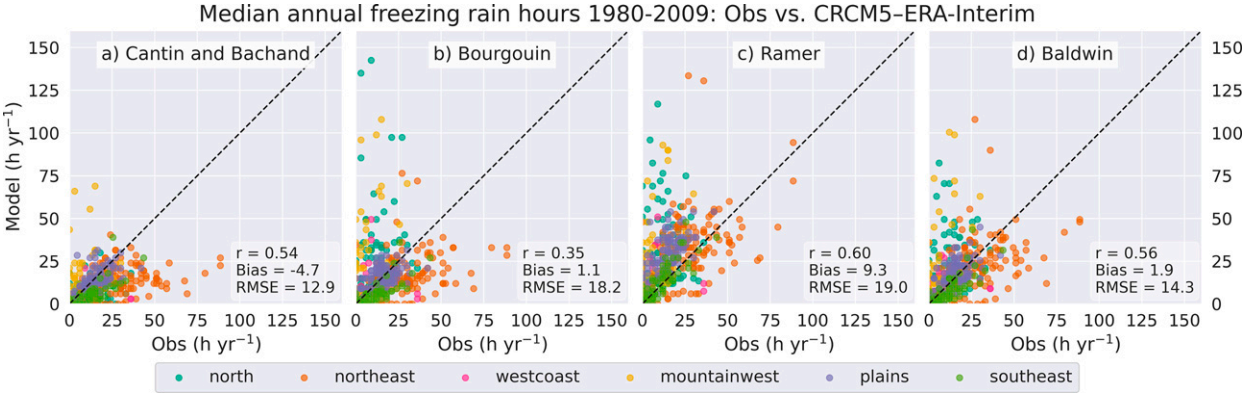

FIG. 4. Median annual hours of freezing rain (1980–2009) in the CRCM5–ERA-Interim simulation compared with observations at 579 surface stations in the United States and Canada for the (a) Cantin and Bachand, (b) Bourgouin, (c) Ramer, and (d) Baldwin algorithms. Bukovsky-region groups for each station are color coded as in Fig. 1. Also displayed are the Pearson correlation coefficient r, mean bias, and RMSE for each method compared with observations among all stations.

TABLE 2. Evaluation of simulated median annual hours of freezing rain compared with observations for six regions as well as for the entire dataset of 579 stations. Pearson correlation coefficients between observed and simulated median annual hours for each method and region are displayed, with mean biases (h yr^{-1}) in parentheses. Algorithms with the best (most positive) correlation and smallest mean bias for each region are indicated by the boldface text.

	North	Northeast	Southeast	Plains	Mountain West	West Coast	All stations
Cantin and Bachand	**0.18 (−6.6)**	0.29 (−14.4)	0.82 (−3.1)	0.69 (−1.9)	0.49 **(6.8)**	0.30 **(−1.8)**	0.54 (−4.7)
Bourgouin	0.09 (19.1)	0.36 (−11.3)	0.82 (−2.8)	0.63 **(0.4)**	0.63 (17.3)	0.22 (3.8)	0.35 **(1.1)**
Ramer (RH_{crit} = 80%)	−0.02 (21.4)	0.41 (5.5)	0.90 **(1.0)**	**0.72** (12.7)	0.60 (19.1)	0.24 (5.4)	**0.60** (9.3)
Ramer (RH_{crit} = 90%)	−0.11 (44.6)	**0.43** (27.9)	**0.91** (4.0)	0.62 (27.4)	0.43 (42.7)	**0.58** (12.9)	0.55 (24.5)
Baldwin (T_{DD} < 6°C)	0.12 (12.7)	0.37 (−4.9)	0.85 (−2.0)	0.71 (1.7)	**0.65** (13.3)	0.30 (2.5)	0.56 (1.9)
Baldwin (T_{DD} < 2°/4°/6°C)	0.06 (21.2)	0.36 **(0.3)**	0.87 (−1.3)	0.49 (8.0)	0.50 (33.2)	0.46 (7.3)	0.42 (8.2)

produce freezing rain frequencies that have the highest correlations with observations, they also have the largest (positive) biases. Bourgouin and Cantin and Bachand have lower biases but also slightly lower correlations with observations and tend to underestimate the maxima in the climatology (Figs. 2a,b).

4. Sensitivity tests

We next examine the sensitivity of the results found in section 3 to various choices that may be necessary when studying freezing rain in climate simulations. We first explore the effect of changing the parameters used for identification of saturated layers in Ramer and Baldwin as well as the sensitivity of each algorithm's results to the choice of minimum precipitation rate above which precipitation type is diagnosed. Finally, we determine the sensitivity of each algorithm to the number of pressure levels available, as climate model output is often only archived on a limited set of levels.

a. Sensitivity to algorithm parameters

First, we compare Ramer using the initial RH_{crit} = 90% threshold for identifying the precipitation generation layer with the 80% threshold found in the Unified Post Processor code. Results are highly sensitive to this choice (Fig. 5). Using RH_{crit} = 90%, freezing rain is identified much more frequently, with broad regions of > 70 h yr^{-1} in the median (Fig. 5a). Such large values are constrained to limited regions of the Pacific Northwest and northern Canada when RH_{crit} = 80% (Fig. 2c). Median freezing rain frequency increases by > 50 h yr^{-1} over much of western, northern, and northeastern North America when RH_{crit} is increased to 90% (Fig. 5c). These increases are attributable to a greater frequency of freezing precipitation formed via the supercooled warm rain process. The more-restrictive RH threshold results in a substantially increased frequency of precipitation forming at warmer vertical levels closer to the surface that are thus identified as freezing rain. Freezing rain frequencies calculated using the RH_{crit} = 90% version of Ramer are better correlated with observations than the RH_{crit} = 80% version in three of the six examined regions (Table 2). However, the mean bias increases with the RH_{crit} = 90% version in all regions, nearly tripling in magnitude compared with the RH_{crit} = 80% version among all stations (24.5 h yr^{-1} as compared with 9.3 h yr^{-1}).

Next, we compare the two thresholds for saturation using Baldwin. Similar to the results for Ramer, a progressively increasing 2°–4°–6°C threshold (Fig. 5b) produces more freezing rain than a fixed T_{DD} < 6°C threshold (Fig. 2d). While the differences are smaller than those for the two variants of Ramer, increases of > 50 h yr^{-1} occur over portions of the Rocky Mountains in the United States and Canada (Fig. 5d). The progressively increasing T_{DD} threshold may identify the precipitation generation layer at relatively warm near-surface layers with T_{DD} < 2°C, ignoring colder layers at higher altitudes with slightly larger T_{DD}. This can result in the identification of freezing rain formed through the supercooled warm rain process in situations where the T_{DD} < 6°C version may identify snow. Regional skill differences between the two versions of Baldwin are similar to the differences between the two versions of Ramer, though bias differences are smaller in most regions than those between the two Ramer variants (Table 2).

Among all United States and Canadian stations, the RH_{crit} = 80% version of Ramer and T_{DD} < 6°C version of Baldwin performed better [lower bias and root-mean-square error (RMSE); higher correlation coefficients] than the RH_{crit} = 90% version of Ramer and the T_{DD} < 2°–4°–6°C version of Baldwin (Table 2). Given their improved overall skill, and in particular their much lower biases, analysis for the remainder of this paper will be restricted to Ramer with RH_{crit} = 80% and Baldwin with a fixed T_{DD} < 6°C threshold.

b. Sensitivity to precipitation thresholds

The overproduction of very light precipitation in climate simulations has led past studies to set a minimum precipitation threshold above which to examine precipitation types (e.g., Lambert and Hansen 2011; Bresson et al. 2017; Matte et al. 2019; Kämäräinen et al. 2018). Because freezing rain is typically of light intensity (Cortinas 2000), the simulated frequency of freezing precipitation may be highly sensitive to changes in the minimum threshold chosen. Studies with the CRCM5 have commonly employed a threshold of 1 mm day^{-1} (Bresson et al. 2017; Matte et al. 2019; St-Pierre et al. 2019), as did Lambert and Hansen (2011) using a GCM. Kämäräinen et al. (2018) tuned the precipitation threshold separately for each GCM-RCM combination they studied in the European CORDEX (EURO-CORDEX; Jacob et al. 2014) ensemble, such that the number of European freezing rain events identified during the 1971–2000 baseline period in the historical simulations most closely matched those identified in ERA-Interim. This led to a broad range of seven minimum precipitation rates, from 1.48 mm day^{-1} for the RCA4 simulation driven by

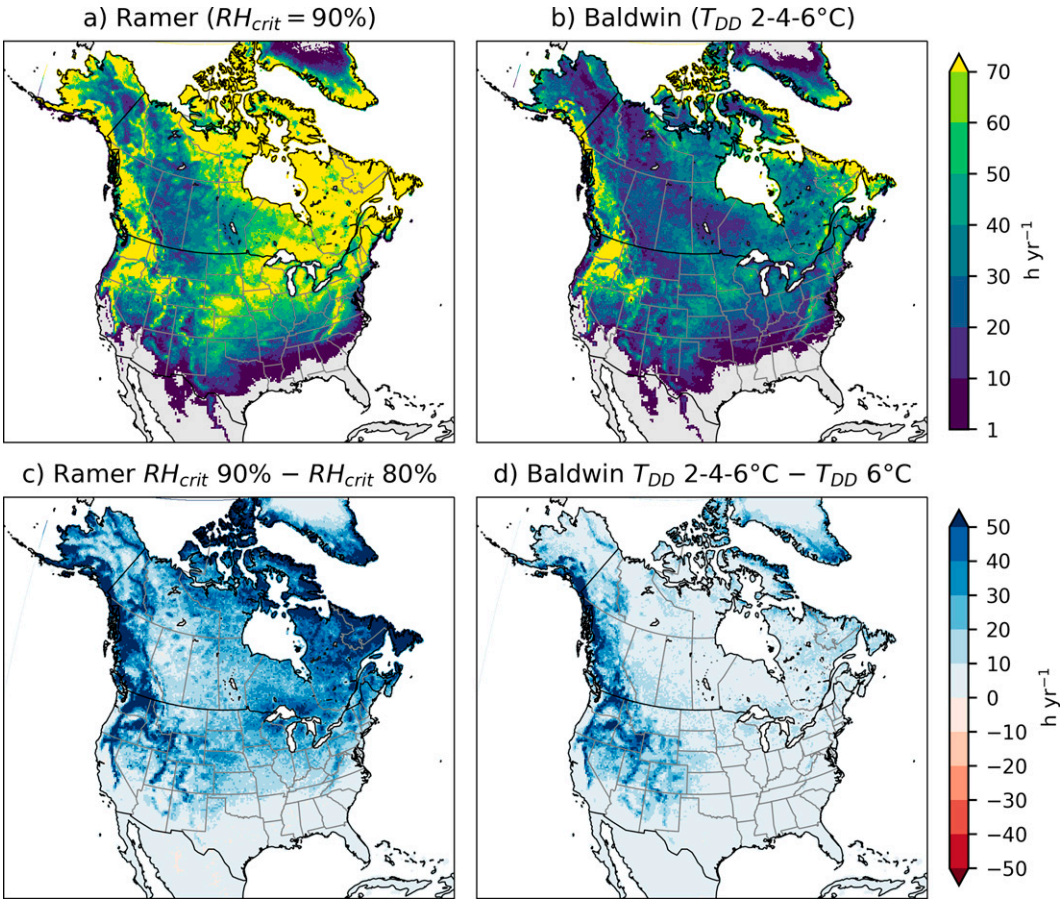

FIG. 5. Median annual hours of freezing rain (1980–2009) in the CRCM5–ERA-Interim simulation calculated using (a) Ramer using a threshold of RH_{crit} = 90% for saturation and (b) Baldwin using an iterative T_{DD} threshold of 2°, 4°, and 6°C. Also shown is the difference (h yr^{-1}) in the median annual hours of freezing rain identified using (c) the Ramer algorithm with a value of RH_{crit} = 90% for saturation minus that obtained using RH_{crit} = 80% and (d) the difference between values obtained from the Baldwin algorithm using a progressively increasing threshold of 2°, 4°, and 6°C minus those obtained using a fixed $T_{DD} > 6$°C threshold.

IPSL-CM5A-MR to 11.4 mm day^{-1} the MPI Regional Model 2009 (REMO2009) driven by MPI-ESM-LR.

We have tested our results using several precipitation thresholds but focus here on the difference between using 1 and 2 mm day^{-1} for brevity (Fig. 6). As expected, freezing precipitation frequency decreases as the threshold increases and the frequency of precipitation decreases, though the magnitude of decrease is highly dependent on the algorithm. Ramer, with the highest initial frequency of freezing rain (Fig. 2b), sees the largest absolute and relative decreases in freezing rain frequency (domain average of −48%) when the threshold is doubled, and decreases of over 20 h yr^{-1}

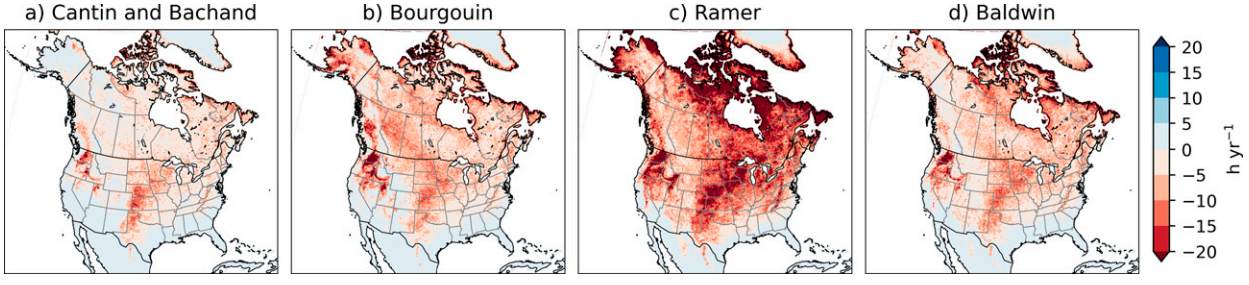

FIG. 6. Difference between the 1980–2009 median annual hours of freezing rain identified from the CRCM5–ERA-Interim simulation using a minimum precipitation threshold of 2 and 1 mm day^{-1} (frequencies obtained using 2 mm day^{-1} minus frequencies obtained using 1 mm day^{-1}) for the (a) Cantin and Bachand, (b) Bourgouin, (c) Ramer, and (d) Baldwin algorithms.

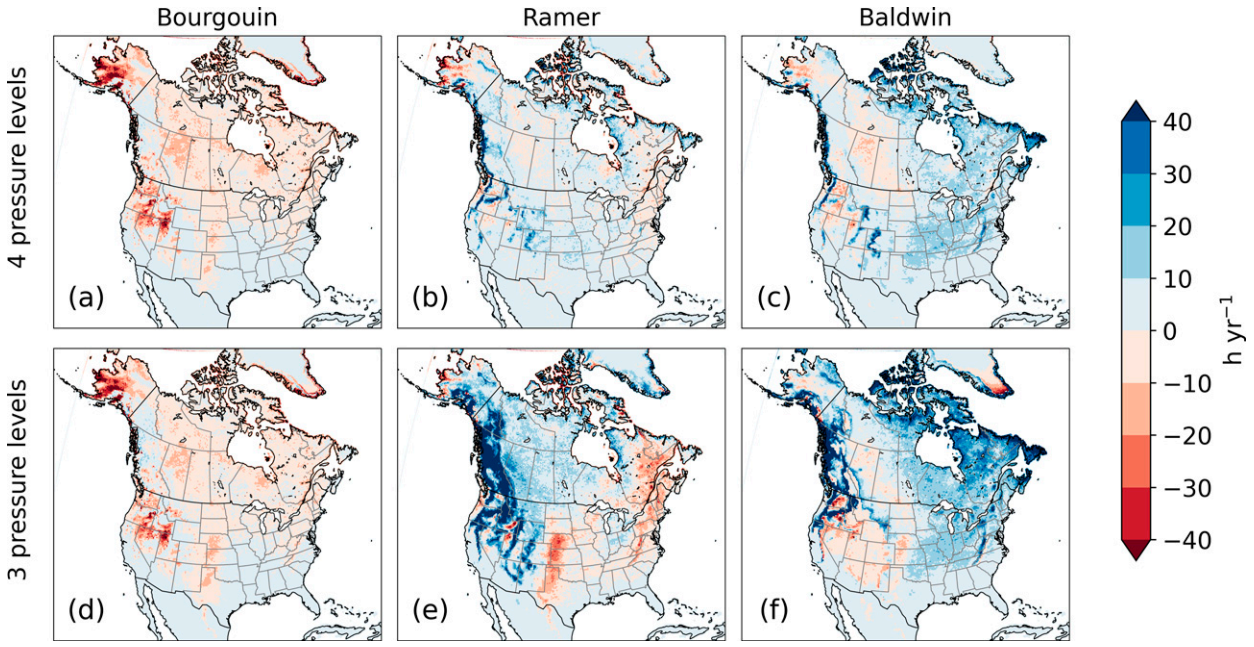

FIG. 7. Effect of the number of pressure levels used on the 1980–2009 median annual hours of freezing rain identified in the CRCM5–ERA-Interim simulation. The medians presented in Fig. 2 using 27 pressure levels are subtracted from those calculated using only (top) 4 levels (500, 700, 850, and 1000 hPa) or (bottom) 3 levels (500, 850, and 1000 hPa) for the (a),(d) Bourgouin; (b),(e) Ramer; and (c),(f) Baldwin algorithms.

are common (Fig. 6c). Bourgouin and Baldwin show smaller decreases (domain-averaged decreases of 39% and 37%, respectively) and are generally similar to each other (Figs. 6b,d). Cantin and Bachand shows the smallest sensitivity to precipitation threshold (domain-averaged decrease of 35%, Fig. 6a).

The decreases in Ramer and Baldwin are largely due to a substantial reduction in freezing rain formed via the supercooled warm rain process, which shows a domain-averaged decrease of 61% for Ramer and 74% for Baldwin compared with 25% and 24%, respectively, for freezing rain formed through the melting process. This is consistent with past findings that the supercooled warm rain process is most commonly associated with lighter freezing drizzle (Huffman and Norman 1988; Rauber et al. 2000).

c. *Sensitivity to number of vertical levels*

To identify freezing rain formed through the melting process, data must be available on a sufficient number of vertical levels to accurately represent the warm layer aloft and near-surface cold layer. This is often a challenge as subdaily climate model output is often archived on only a few pressure levels. To explore the sensitivity of algorithms to the number of levels used, we diagnose precipitation types from the CRCM5–ERA-Interim simulation using data on the surface and three pressure levels: 500, 850, and 1000 hPa, as in Jeong et al. (2019). We also test the effect of the addition of 700 hPa. These levels are chosen both because they are commonly archived and given their importance for reproducing the warm and cold layers. We then compare these results with those derived from the full set of 27 vertical levels available (Fig. 7). Cantin and Bachand is omitted as it only uses data at 500, 700, 850, and 1000 hPa and is therefore either unchanged or incalculable for the two scenarios presented here.

First, Bourgouin (Figs. 7a,d) is the least sensitive method to the number of pressure levels, with a general slight decrease in freezing rain frequency using three or four instead of all 27 levels, and an additional small reduction in frequency when going from four to three levels. The largest decreases occur in the higher-terrain regions of Alaska and the northwestern United States. These results are consistent with Jeong et al. (2019), who calculated freezing precipitation frequency using Bourgouin with 3, 6, and 11 vertical levels in NARR and found that results were generally similar for the three sets of levels.

Unlike Bourgouin, Ramer exhibits an increase in freezing rain frequency at most grid points when moving from 27 to 4 levels, with the largest increases occurring over the Rocky Mountains (Fig. 7b). More substantial changes occur when reducing to three levels, with large increases in freezing rain frequency over the mountainous western United States and Canada (Fig. 7e). Median values of > 200 h yr^{-1} occur near the coast of British Columbia using only three levels. The largest decreases occur over the western Great Plains (from Texas northward to Nebraska), with slight decreases also present over much of the northeastern United States and eastern Canada (Fig. 7e).

Baldwin is nearly as sensitive to the number of pressure levels as Ramer, though the spatial structure of the change is different particularly for the three-level calculation (Figs. 7c,f). Freezing rain frequency increases as the number of levels decrease at most locations, with the largest increases occurring

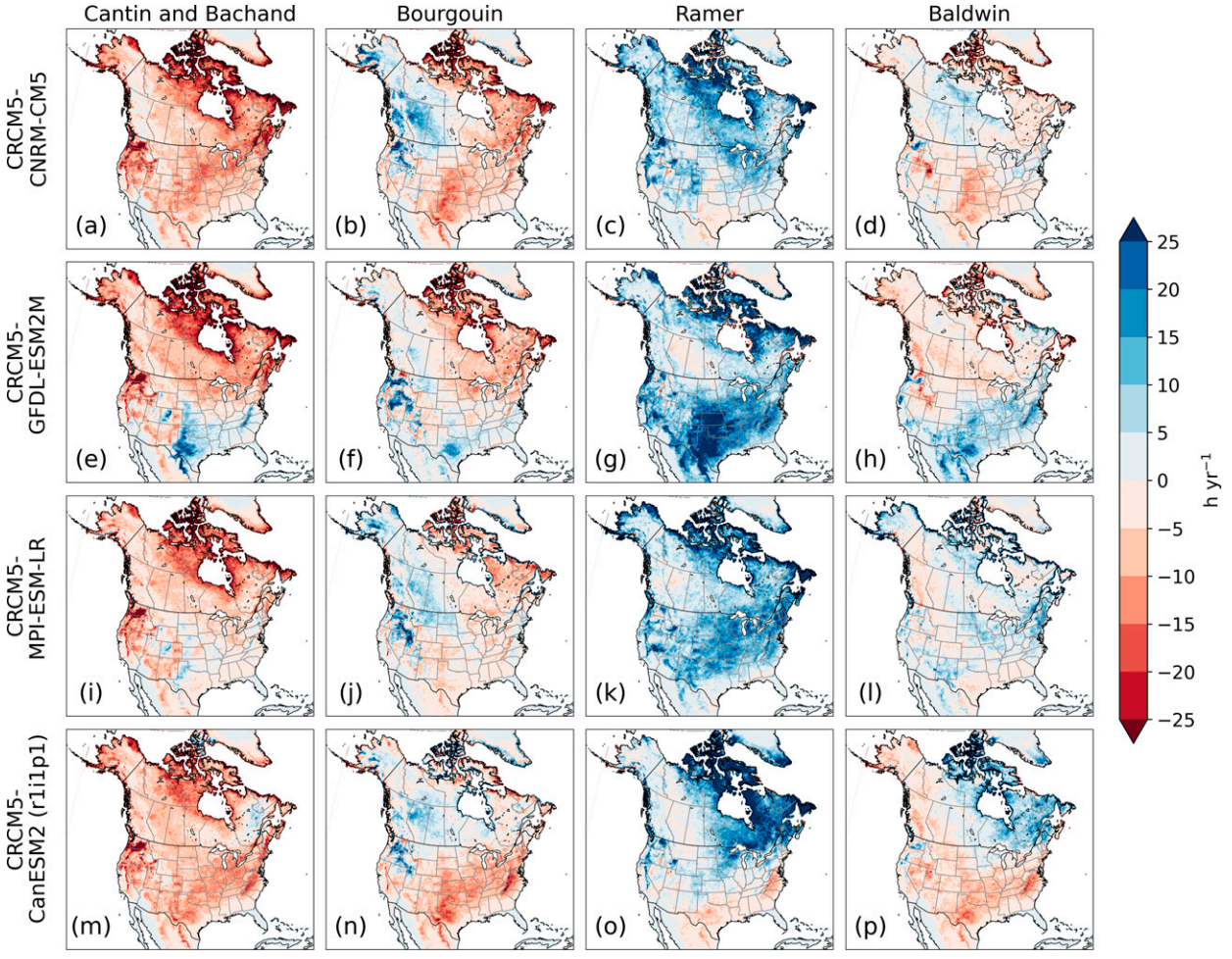

FIG. 8. Deviation (h) from the overall multisimulation/multialgorithm mean of the median annual hours of freezing rain (1980–2009) for each simulation–algorithm combination. Algorithms displayed are (left) Cantin and Bachand, (left center) Bourgouin, (right center) Ramer, and (right) Baldwin for CRCM5 simulations driven by (a)–(d) CNRM-CM5, (e)–(h) GFDL-ESM2M, (i)–(l) MPI-ESM-LR, and (m)–(p) CanESM2. Only the first member of the CRCM5-CanESM2 ensemble is displayed, because spatial patterns of the other four members are very similar.

over the higher terrain of western Canada and the northwestern United States as in Ramer. Unlike Ramer, large increases occur over Newfoundland, Labrador, and northern Québec for the three-level calculations (Fig. 7c).

The large increases, especially over mountainous regions, in Ramer and Baldwin can be partly explained by the ability of these algorithms to identify freezing precipitation formed through the supercooled warm rain process. At locations where surface pressure is < 1000 hPa, the three-level calculations rely only on data at 500 and 850 hPa and the surface. In situations where 500 hPa is unsaturated, the next possible precipitation generation level is 850 hPa. If the 850-hPa temperature is too warm for ice nucleation (according to the algorithm criteria presented in section 2), precipitation will be identified as liquid, which will then be diagnosed as freezing rain if surface temperature < 0°C. With 27 vertical levels, saturation may have occurred at colder temperatures at a level between 500 and 850 hPa and instead may have been diagnosed as snow.

5. Comparison with other sources of uncertainty

We now examine the variability in the climatology of freezing rain produced by the CRCM5 driven by different GCMs. Additionally, we analyze an ensemble of CanESM2-driven simulations to estimate natural variability. These analyses will allow us to place the uncertainty associated with algorithm selection in the context of these other key sources of uncertainty in climate simulations of freezing rain.

a. Uncertainty related to choice of driving GCM

The choice of driving GCM can have a large influence on the RCM-simulated climate, with variability sometimes larger than that between different RCMs for certain variables (e.g., Christensen and Kjellström 2020). This is due to several factors, notably that each GCM has its own biases that influence the RCM through the prognostic variables driven at the boundaries. Here, we are interested in how uncertainty

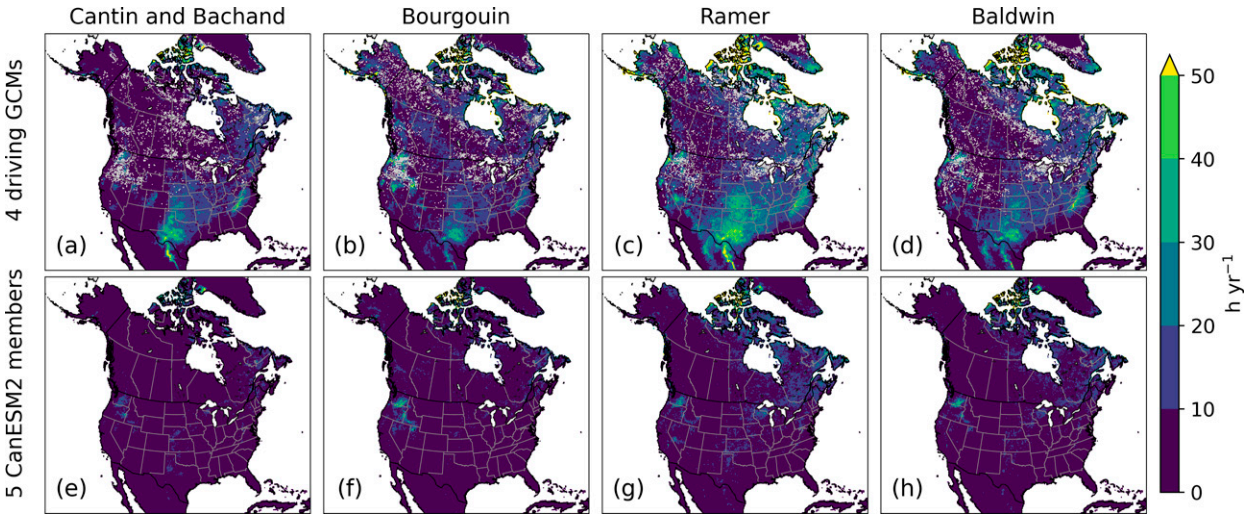

FIG. 9. Range (maximum − minimum) of values of median annual hours of freezing rain (1980–2009) obtained from CRCM5 driven by (top) four different GCMs or (bottom) the five CanESM2 ensemble members for the (a),(e) Cantin and Bachand; (b),(f) Bourgouin; (c),(g) Ramer; and (d),(h) Baldwin algorithms. Gray-shaded grid points in (a)–(d) are areas where the range of values among the five CanESM2-driven ensemble members is greater than the range of the four driving GCMs, as discussed in section 5b.

associated with choice of precipitation-type algorithm compares with that associated with choice of driving GCM. Figure 8 presents the median annual hours of freezing rain identified using the four algorithms and the four driving GCMs listed in Table 1 minus the overall multialgorithm/multisimulation mean. Examination of a given row allows for a visual analysis of the effect of algorithm selection for a given driving GCM, while examination of a single column allows for a visual analysis of the effect of driving GCM for a given algorithm.

Consistent with the previously presented results, the Ramer algorithm produces the most freezing rain for each simulation (Figs. 8c,g,k,o), while Cantin and Bachand produces the least (Figs. 8a,e,i,m). However, large regional differences are present depending on driving GCM. The GFDL-ESM2M-driven simulation produces the most freezing rain over the southern United States regardless of algorithm (Figs. 8e–h). The simulations driven by CNRM-CM5 (Figs. 8a–d) and CanESM2 (Figs. 8m–p) produce the least freezing rain overall, while the MPI-ESM-LR-driven simulation is closest to the overall mean in most regions (Figs. 8i–l).

The largest variability between the simulations is found over the southern United States (Figs. 9a–d), where CRCM5–GFDL-ESM2M produces more frequent freezing rain than the other simulations (Figs. 8e–h). The largest interalgorithm variability is found in the Canadian Arctic, Newfoundland, Alaska, and the Pacific Northwest regardless of driving GCM (Fig. 10). The spatial pattern of interalgorithm variability is similar between driving GCMs with the exception of the CRCM5–GFDL-ESM2M, which has much larger variability over the U.S. southern Great Plains (Fig. 10b). Comparison of Figs. 9 and 10 highlights that the variability in the simulated climatology of freezing rain among different algorithms is at least of similar magnitude and in most regions is larger than variability between driving GCMs.

Differences among algorithms and driving GCMs can also be visualized using Taylor (2001) diagrams, displaying the skill of each algorithm–simulation combination at representing the observed climatology of freezing rain. The results for a given algorithm are closely clustered (Fig. 11), highlighting the importance of algorithm selection on the resultant freezing

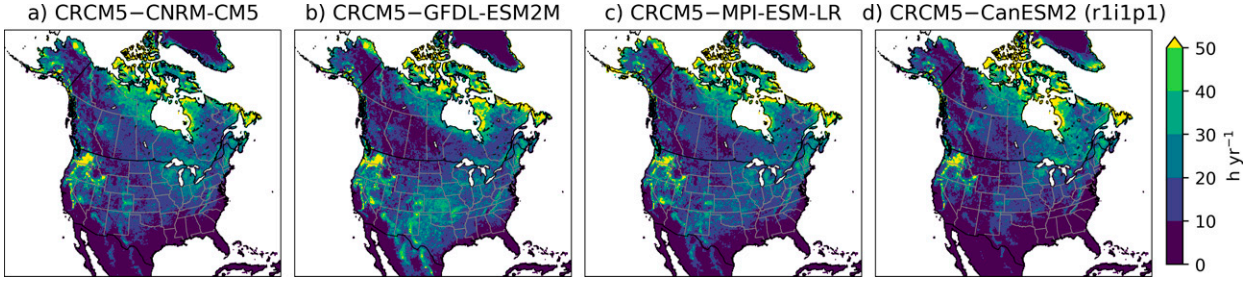

FIG. 10. Similar to Fig. 9, but for the range (maximum − minimum) of values of median annual hours of freezing rain (1980–2009) obtained using the four different algorithms discussed in the text for CRCM5 driven by (a) CNRM-CM5, (b) GFDL-ESM2M, (c) MPI-ESM-LR, and (d) CanESM2 (r1i1p1).

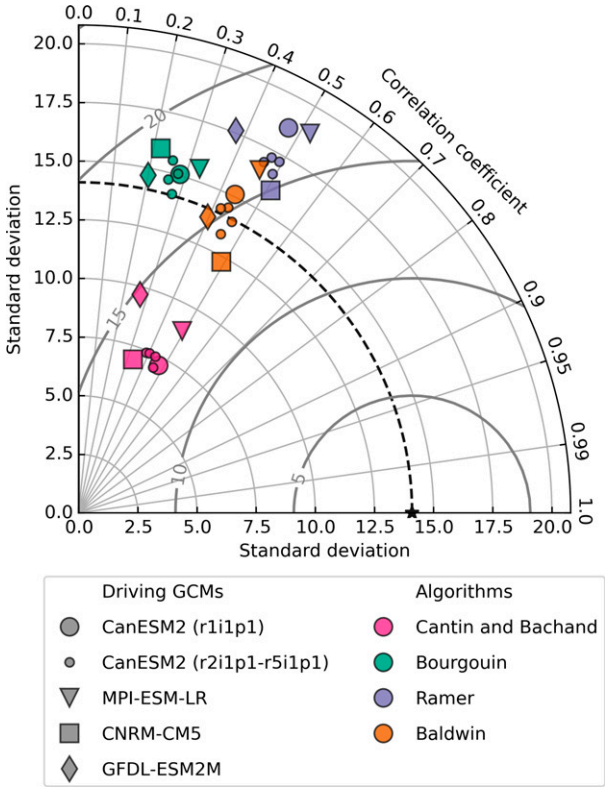

FIG. 11. Taylor (2001) diagram comparing observed and simulated values of the 1980–2009 median annual hours of freezing rain. CRCM5 driven by different GCMs are identified by the symbols described in the legend, and the precipitation-type algorithm is indicated by the color of the points. Pearson correlation coefficients between observed and simulated frequencies are displayed by the position of the points relative to the gray rays. The standard deviation of the observations is indicated by the black dashed curve, and that of the simulations is shown by the position of the points relative to the thin gray curves. The RMSE of the simulated climatologies is shown by the position of the points relative to the thick gray curves.

rain climatology. For example, Cantin and Bachand underestimates spatial variability, with a lower standard deviation for all driving GCMs compared with observations, while Ramer overestimates spatial variability.

Separating the Taylor (2001) diagrams by region produces similar results but highlights the strong geographic dependence of model skill (Fig. 12). Though results for a given algorithm still tend to cluster in most regions, skill in some regions is more dependent on driving GCM. For example, in the West Coast (Fig. 12f), results for Bourgouin, Ramer, and Baldwin are clustered together by driving GCM, while Cantin and Bachand results for all driving GCMs are clustered. In the Northeast (Fig. 12b), all combinations of driving GCM and algorithm underestimate the observed spatial variability. Conversely, spatial variability is overestimated by all simulations and algorithms in the Mountain West (Fig. 12e). In most regions, points are aligned along a line of roughly constant correlation coefficient, regardless of simulation. This may be related to structural biases in the CRCM5 discussed in section 3 or the model representation of key terrain features in certain regions. For example, spatial correlations for all combinations of algorithms and driving GCMs in the North and West Coast (Figs. 12a,f) are low, while in the southeast $r > 0.7$ for all combinations (Fig. 12c).

To quantify each simulation's ability to reproduce the climatology of freezing rain, we take the multialgorithm mean of the median annual hours of freezing rain at each station and compare with observations (Table 3). Among all stations, the CRCM5–MPI-ESM-LR simulation has the strongest correlation with observations, while the CRCM5–GFDL-ESM2M has the worst correlation but the smallest mean bias. The poor skill of the CRCM5–GFDL-ESM2M is partly due to its overestimation of freezing rain over portions of the United States as previously discussed, namely the Plains region where the CRCM5–GFDL-ESM2M has the lowest correlation ($r = 0.34$) and highest bias (6.0 h yr^{-1}) among the simulations. Scores are generally similar between simulations, though the CRCM5–MPI-ESM-LR has the best scores in the most regions, followed by CRCM5–CNRM-CM5 and then CRCM5–CanESM2.

b. *Natural variability*

Five CRCM5 simulations driven by different CanESM2 members are available. Though a small number of members, this allows us to gain some insight into the magnitude of natural variability relative to the variability between algorithms and between driving GCMs. Spread in the medians among the different CRCM5–CanESM2 members is generally very small, <10 h yr^{-1} at 80% (Ramer) to 95% (Cantin and Bachand) of land grid points (Figs. 9e–h). The largest spread is found over the simulated maximum in freezing rain frequency in the Pacific Northwest and the Canadian Arctic Archipelago.

Overall, uncertainties associated with the choice of driving GCM (Figs. 9a–d) and algorithm (Fig. 10) are larger than that associated with natural variability (Figs. 9e–h). Grid points where natural variability is larger than variability between driving GCMs in Fig. 9 are shaded gray. This is the case primarily at locations where freezing rain is relatively infrequent and where variability among driving GCMs is low as well as over the Pacific Northwest, where both simulated freezing rain frequency and variability among driving GCMs are high. The range of the medians calculated for the four driving GCMs for a given algorithm is larger than that between the five CanESM2-driven simulations at most locations, ranging from 90% of land grid points with Ramer to 93% for Cantin and Bachand. The values of the five CanESM2-driven members are also closely clustered together in the Taylor diagrams (Figs. 11, 12), highlighting that natural variability is smaller than variability associated with either driving GCM or algorithm.

6. Concluding discussion and recommendations

In this paper, we have applied four precipitation-type algorithms to an ensemble of simulations of the fifth-generation Canadian Regional Climate Model (CRCM5) over North America at 0.22° grid spacing. We have examined the sensitivity of the simulated freezing rain climatology to the choices

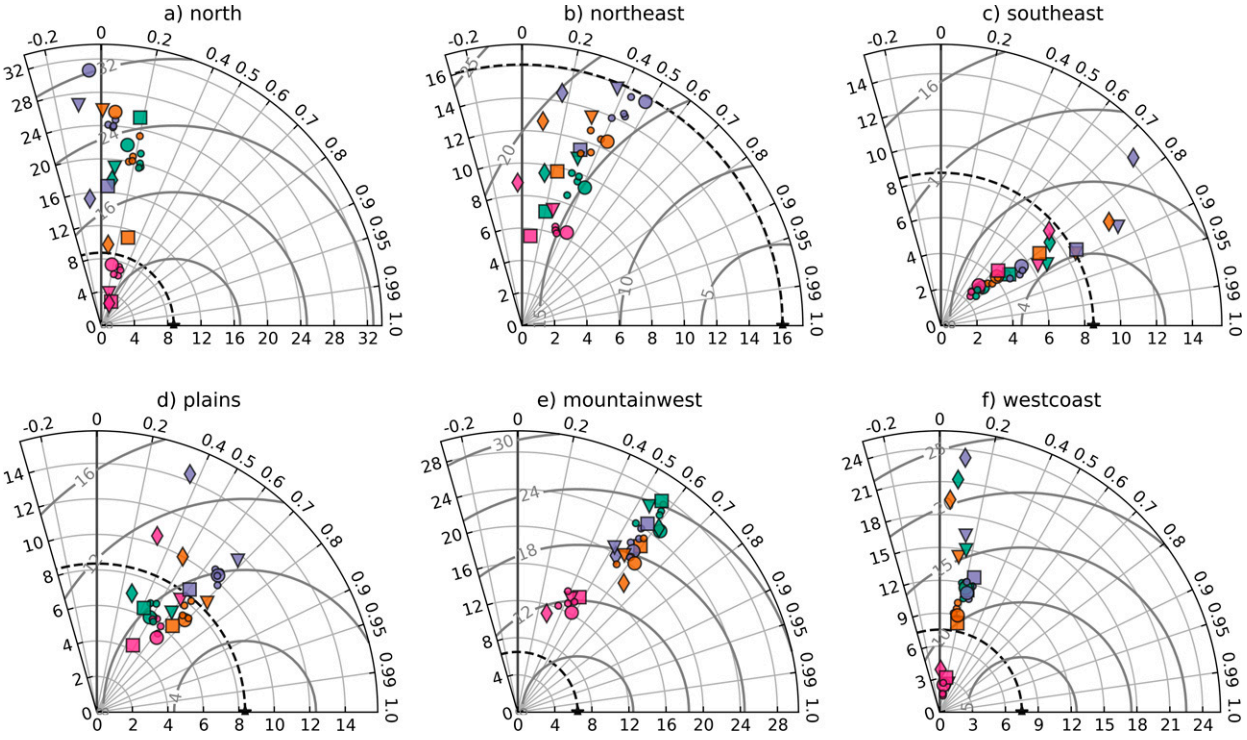

FIG. 12. Taylor (2001) diagrams as in Fig. 11, but for each group of Bukovsky (2011) regions displayed in Fig. 1: (a) North, (b) Northeast, (c) Southeast, (d) Plains, (e) Mountain West, and (f) West Coast.

of algorithm, algorithm parameters, minimum precipitation thresholds, the number of pressure levels used, and the GCM used to drive the CRCM5. We have demonstrated that the results obtained can vary greatly between algorithms, but also between details of the model and algorithm configurations chosen.

No perfect precipitation-type algorithm exists, with each method having particular biases (e.g., Reeves et al. 2014). Applying the algorithms to the CRCM5 driven by ERA-Interim, we have found that the Ramer (1993) and Baldwin and Contorno (1993) algorithms produced freezing rain frequencies that were better correlated with observations compared with the Cantin and Bachand (1993) and Bourgouin (2000) techniques (Table 2). However, the Ramer and Baldwin methods are highly sensitive to the number of pressure levels available, the minimum precipitation rate threshold applied, and to saturation thresholds for identifying the precipitation generation layer. This sensitivity stems largely from the ability of these methods to identify freezing precipitation formed through the supercooled warm rain process. Use of higher relative humidity and lower dewpoint depression thresholds led to an overestimation of freezing rain frequency compared with observations (Table 2). We therefore recommend the use of lower thresholds for application to climate model output to avoid identifying too-frequent freezing rain formed through the supercooled warm rain process.

The sensitivity of the Baldwin and especially Ramer methods to the number of pressure levels used suggests these methods may not be appropriate for climate studies that often rely on a limited number of archived vertical levels. The impact is particularly large when reducing from 27 to three levels (500, 850, and 1000 hPa), with very large increases in freezing rain frequencies over some regions. The addition of 700 hPa produces a result much closer to that for 27 levels, suggesting these four pressure levels may be a minimum required to obtain reasonable results with Ramer and Baldwin. Though results obtained using the Bourgouin algorithm are generally slightly less correlated with observed freezing rain frequencies than Ramer or Baldwin, Bourgouin is more robust to the

TABLE 3. As in Table 2, except displaying Pearson correlation coefficients between the observed median annual hours of freezing rain and the four-algorithm mean for each simulation and region, with mean biases (h yr^{-1}) in parentheses. Simulations with the best (most positive) correlation and smallest mean bias for each region are indicated by the boldface text.)

	North	Northeast	Southeast	Plains	Mountain West	West Coast	All stations
CRCM5–CNRM-CM5	**0.16** (5.3)	0.41 (−11.9)	0.83 (−2.1)	0.64 **(−2.2)**	0.59 (15.7)	**0.24 (1.8)**	0.46 (−2.2)
CRCM5–GFDL-ESM2M	0.02 **(1.0)**	0.20 (−9.5)	0.81 (2.9)	0.34 (6.0)	0.61 (16.0)	0.08 (7.3)	0.38 **(1.2)**
CRCM5–MPI-ESM-LR	−0.07 (9.4)	0.51 **(−7.1)**	**0.89 (−0.0)**	**0.70** (4.1)	0.58 (14.3)	0.16 (4.6)	**0.51** (1.5)
CRCM5–CanESM2 (r1i1p1)	0.01 (8.3)	**0.63** (−12.9)	0.78 (−4.0)	0.68 (−2.9)	**0.62 (12.5)**	0.21 (2.0)	0.46 (−2.9)

number of pressure levels used and may thus be better suited for situations where few vertical levels are available.

Freezing precipitation formed through the supercooled warm rain process is in many cases freezing drizzle rather than freezing rain (Rauber et al. 2000). In observations, freezing drizzle is typically reported separately from freezing rain and can represent a large percentage of freezing precipitation observations over some regions (e.g., Cortinas et al. 2004). This results in more freezing precipitation production in these algorithms than Bourgouin or Cantin and Bachand, and in some cases a large overestimation when compared with the observed climatology of freezing rain. For future climate studies, differences in the conditions leading to freezing drizzle and freezing rain should be addressed. Freezing drizzle typically occurs with a shallow near-surface saturated layer, often following the passage of an Arctic cold front (Rauber et al. 2001). Freezing rain is instead typically associated with midlatitude cyclones. Changes to the frequency of the meteorological patterns leading to the two precipitation types may result in different responses in a warming climate.

The Cantin and Bachand method requires the least model output, with geopotential height on only three levels necessary. Because this method does not take advantage of the detailed near-surface information available from the pressure levels used in the other methods, Cantin and Bachand underestimates the observed maximum freezing rain frequencies in northeastern North America. However, this technique does demonstrate comparable skill to the other methods over some regions (Table 2). Cantin and Bachand may therefore be a good candidate for simulations with limited thermodynamic model output. A key limitation of this method is that it is not appropriate for higher terrain locations due to its use of pressure levels that may be below ground at these locations.

Algorithm skill is also dependent on model temperature and precipitation biases. For example, all algorithms in the CRCM5–ERA-Interim simulation performed relatively poorly over northern Canada and Alaska and relatively well over the southeastern United States (Table 2), likely related to structural biases in the CRCM5. Our comparison of the results obtained from the four CRCM5 simulations driven by different GCMs demonstrates that the uncertainty associated with choice of driving GCM is large for the recent past climate, with differences between the various simulations exceeding 40 h yr^{-1} in some regions. Conversely, the range of values obtained among CRCM5 simulations driven by five different CanESM2 members, a measure of natural variability, is comparatively small. The range of values identified among the four algorithms for a given GCM is similar to or larger than that among the four CRCM5 simulations driven by different GCMs in many regions. These comparisons highlight both the importance of the uncertainty related to choice of driving GCM but also that of the choice of precipitation-type algorithm. Future studies on freezing rain in climate model output should therefore consider an examination of multiple algorithms and multiple driving GCMs to better encompass this variability.

Our analysis was limited to four commonly used precipitation-type algorithms developed for NWP models. Newer, more sophisticated techniques (e.g., Reeves 2016) including statistical/machine learning methods (e.g., Scheuerer et al. 2017) merit further study. As with the older methods examined here, a key limitation is the data available from model output. With machine learning techniques, a sufficient training dataset against which to compare is a key challenge, as even reanalysis-driven simulations are not expected to precisely reproduce individual observed freezing rain events. Given the regional differences in skill among the algorithms, future studies may wish to weight algorithms based on their skill at identifying freezing rain in the past climate over their region of interest.

Despite their differing results and sensitivities, each of the four algorithms was capable of reproducing the general pattern of the observed freezing rain climatology over North America using 17–27 vertical levels at 0.22° grid spacing, lending confidence to their use in climate studies. The next step is therefore to determine the extent to which projected changes to this climatology vary among algorithms. Given the dependence of Ramer and Baldwin on the vertical moisture profile, changes in atmospheric moisture content may result in different changes to freezing rain events for these algorithms than for the Bourgouin or Cantin and Bachand methods that have no moisture criteria. These will be examined in a forthcoming study and will allow for an overall quantification and comparison of the different sources of uncertainty associated with climate model projections of freezing rain.

Acknowledgments. This research was funded through a Mitacs Accelerate grant as well as the first author's Fonds de recherche du Québec–Nature et technologies (FRQNT) postdoctoral research fellowship. The CRCM5 data were generated and supplied by Ouranos. The CRCM5 was developed at the Université du Québec à Montréal (UQAM) in collaboration with Environment and Climate Change Canada (ECCC). The CRCM5 computations were completed on the Guillimin supercomputer at McGill University, managed by Calcul Québec and Compute Canada. Operation of the supercomputer is funded by the Canada Foundation for Innovation (CFI), the Ministère de l'Économie et de l'Innovation du Québec (MEI), and the FRQNT. The authors thank Dominic Matte, Michael Baldwin, and Sébastien Marinier for their assistance with code for the precipitation-type algorithms. Thanks are given to Sébastien Biner for his assistance with the Bukovsky (2011) region masks.

Data availability statement. The CRCM5 model output analyzed here can be obtained by contacting scenarios@ouranos.ca. The observational dataset compared with model output and described in McCray et al. (2019) was derived from the Integrated Surface Database (Smith et al. 2011, https://www.ncei.noaa.gov/products/land-based-station/integrated-surface-database).

REFERENCES

Arora, V. K., and Coauthors, 2011: Carbon emission limits required to satisfy future representative concentration pathways of

greenhouse gases. *Geophys. Res. Lett.*, **38**, L05805, https://doi.org/10.1029/2010GL046270.

Baldwin, M. E., and S. Contorno, 1993: Development of a weather-type prediction system for NMC's mesoscale Eta model. *13th Conf. on Weather Analysis and Forecasting*, Vienna, VA, Amer. Meteor. Soc., 86–87.

Benjamin, S. G., J. M. Brown, and T. G. Smirnova, 2016: Explicit precipitation-type diagnosis from a model using a mixed-phase bulk cloud–precipitation microphysics parameterization. *Wea. Forecasting*, **31**, 609–619, https://doi.org/10.1175/WAF-D-15-0136.1.

Bernstein, B. C., 2000: Regional and local influences on freezing drizzle, freezing rain, and ice pellet events. *Wea. Forecasting*, **15**, 485–508, https://doi.org/10.1175/1520-0434(2000)015<0485:RALIOF>2.0.CO;2.

Birk, K., E. Lenning, K. Donofrio, and M. T. Friedlein, 2021: A revised Bourgouin precipitation-type algorithm. *Wea. Forecasting*, **36**, 425–438, https://doi.org/10.1175/WAF-D-20-0118.1.

Bourgouin, P., 2000: A method to determine precipitation types. *Wea. Forecasting*, **15**, 583–592, https://doi.org/10.1175/1520-0434(2000)015<0583:AMTDPT>2.0.CO;2.

Bresson, E., R. Laprise, D. Paquin, J. M. Thériault, and R. de Elía, 2017: Evaluating the ability of CRCM5 to simulate mixed precipitation. *Atmos.–Ocean*, **55**, 79–93, https://doi.org/10.1080/07055900.2017.1310084.

Brooks, C. F., 1920: The nature of sleet and how it is formed. *Mon. Wea. Rev.*, **48**, 69–72, https://doi.org/10.1175/1520-0493(1920)48<69b:TNOSAH>2.0.CO;2.

Bukovsky, M., 2011: Masks for the Bukovsky regionalization of North America. North American Regional Climate Change Assessment Program, http://www.narccap.ucar.edu/contrib/bukovsky/.

Cantin, A., and D. Bachand, 1993: Synoptic pattern recognition and partial thickness techniques as a tool for precipitation types forecasting associated with a winter storm. Centre Meteorologique du Quebec Tech. Note 93n-002, 9 pp.

Carrera, M. L., J. R. Gyakum, and C. A. Lin, 2009: Observational study of wind channeling within the St. Lawrence River valley. *J. Appl. Meteor. Climatol.*, **48**, 2341–2361, https://doi.org/10.1175/2009JAMC2061.1.

Changnon, S. A., 2003: Characteristics of ice storms in the United States. *J. Appl. Meteor.*, **42**, 630–639, https://doi.org/10.1175/1520-0450(2003)042<0630:COISIT>2.0.CO;2.

Cholette, M., R. Laprise, and J. Thériault, 2015: Perspectives for very high-resolution climate simulations with nested models: Illustration of potential in simulating St. Lawrence River valley channelling winds with the fifth-generation Canadian Regional Climate Model. *Climate*, **3**, 283–307, https://doi.org/10.3390/cli3020283.

Christensen, O. B., and E. Kjellström, 2020: Partitioning uncertainty components of mean climate and climate change in a large ensemble of European regional climate model projections. *Climate Dyn.*, **54**, 4293–4308, https://doi.org/10.1007/s00382-020-05229-y.

Cortinas, J., Jr., 2000: A climatology of freezing rain in the Great Lakes region of North America. *Mon. Wea. Rev.*, **128**, 3574–3588, https://doi.org/10.1175/1520-0493(2001)129<3574:ACOFRI>2.0.CO;2.

——, K. F. Brill, and M. E. Baldwin, 2002: Probabilistic forecasts of precipitation type. *16th Conf. on Probability and Statistics in the Atmospheric Sciences*, Orlando, FL, Amer. Meteor. Soc., 3.13, https://ams.confex.com/ams/pdfpapers/30176.pdf.

——, B. C. Bernstein, C. C. Robbins, and J. W. Strapp, 2004: An analysis of freezing rain, freezing drizzle, and ice pellets across the United States and Canada: 1976–90. *Wea. Forecasting*, **19**, 377–390, https://doi.org/10.1175/1520-0434(2004)019<0377:AAOFRF>2.0.CO;2.

Czys, R. R., R. W. Scott, K. C. Tang, R. W. Przybylinski, and M. E. Sabones, 1996: A physically based, nondimensional parameter for discriminating between locations of freezing rain and ice pellets. *Wea. Forecasting*, **11**, 591–598, https://doi.org/10.1175/1520-0434(1996)011<0591:APBNPF>2.0.CO;2.

Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, https://doi.org/10.1002/qj.828.

Dunne, J. P., and Coauthors, 2012: GFDL's ESM2 global coupled climate–carbon Earth system models. Part I: Physical formulation and baseline simulation characteristics. *J. Climate*, **25**, 6646–6665, https://doi.org/10.1175/JCLI-D-11-00560.1.

Giorgetta, M. A., and Coauthors, 2013: Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the Coupled Model Intercomparison Project phase 5. *J. Adv. Model. Earth Syst.*, **5**, 572–597, https://doi.org/10.1002/jame.20038.

Gyakum, J. R., and P. J. Roebber, 2001: The 1998 ice storm-analysis of a planetary-scale event. *Mon. Wea. Rev.*, **129**, 2983–2997, https://doi.org/10.1175/1520-0493(2001)129<2983:TISAOA>2.0.CO;2.

Henson, W., and R. Stewart, 2007: Severity and return periods of icing events in the Montréal area. *Atmos. Res.*, **84**, 242–249, https://doi.org/10.1016/j.atmosres.2006.08.002.

Huffman, G. J., and G. A. Norman, 1988: The supercooled warm rain process and the specification of freezing precipitation. *Mon. Wea. Rev.*, **116**, 2172–2182, https://doi.org/10.1175/1520-0493(1988)116<2172:TSWRPA>2.0.CO;2.

Jacob, D., and Coauthors, 2014: EURO-CORDEX: New high-resolution climate change projections for European impact research. *Reg. Environ. Change*, **14**, 563–578, https://doi.org/10.1007/s10113-013-0499-2.

Jeong, D. I., L. Sushama, M. J. Vieira, and K. A. Koenig, 2018: Projected changes to extreme ice loads for overhead transmission lines across Canada. *Sustainable Cities Soc.*, **39**, 639–649, https://doi.org/10.1016/j.scs.2018.03.017.

——, A. J. Cannon, and X. Zhang, 2019: Projected changes to extreme freezing precipitation and design ice loads over North America based on a large ensemble of Canadian regional climate model simulations. *Nat. Hazards Earth Syst. Sci.*, **19**, 857–872, https://doi.org/10.5194/nhess-19-857-2019.

Kämäräinen, M., and Coauthors, 2018: Estimates of present-day and future climatologies of freezing rain in Europe based on CORDEX regional climate models. *J. Geophys. Res. Atmos.*, **123**, 13 291–13 304, https://doi.org/10.1029/2018JD029131.

Lambert, S. J., and B. K. Hansen, 2011: Simulated changes in the freezing rain climatology of North America under global warming using a coupled climate model. *Atmos.–Ocean*, **49**, 289–295, https://doi.org/10.1080/07055900.2011.607492.

Martynov, A., R. Laprise, L. Sushama, K. Winger, L. Šeparović, and B. Dugas, 2013: Reanalysis-driven climate simulation over CORDEX North America domain using the Canadian regional climate model, version 5: Model performance evaluation. *Climate Dyn.*, **41**, 2973–3005, https://doi.org/10.1007/s00382-013-1778-9.

Matte, D., J. M. Thériault, and R. Laprise, 2019: Mixed precipitation occurrences over southern Québec, Canada, under

warmer climate conditions using a regional climate model. *Climate Dyn.*, **53**, 1125–1141, https://doi.org/10.1007/s00382-018-4231-2.

McCray, C. D., E. H. Atallah, and J. R. Gyakum, 2019: Long-duration freezing rain events over North America: Regional climatology and thermodynamic evolution. *Wea. Forecasting*, **34**, 665–681, https://doi.org/10.1175/WAF-D-18-0154.1.

Mearns, L., and Coauthors, 2017: The NA-CORDEX dataset, version 1.0. NCAR Climate Data Gateway, accessed 28 March 2022, https://doi.org/10.5065/D6SJ1JCH.

Meisinger, C. L., 1920: The precipitation of sleet and the formation of glaze in the eastern United States, January 20 to 25, 1920, with remarks on forecasting. *Mon. Wea. Rev.*, **48**, 73–80, https://doi.org/10.1175/1520-0493(1920)48<73b:TPOSAT>2.0.CO;2.

Mesinger, F., and Coauthors, 2006: North American Regional Reanalysis. *Bull. Amer. Meteor. Soc.*, **87**, 343–360, https://doi.org/10.1175/BAMS-87-3-343.

Mullens, E. D., and R. McPherson, 2017: A multialgorithm reanalysis-based freezing-precipitation dataset for climate studies in the south-central United States. *J. Appl. Meteor. Climatol.*, **56**, 495–517, https://doi.org/10.1175/JAMC-D-16-0180.1.

Ralph, F. M., and Coauthors, 2005: Improving short-term (0–48 h) cool-season quantitative precipitation forecasting. *Bull. Amer. Meteor. Soc.*, **86**, 1619–1632, https://doi.org/10.1175/BAMS-86-11-1619.

Ramer, J., 1993: An empirical technique for diagnosing precipitation type from model output. *Fifth Int. Conf. on Aviation Weather Systems*, Vienna, VA, Amer. Meteor. Soc., 227–230.

Rauber, R. M., L. S. Olthoff, M. K. Ramamurthy, and K. E. Kunkel, 2000: The relative importance of warm rain and melting processes in freezing precipitation events. *J. Appl. Meteor.*, **39**, 1185–1195, https://doi.org/10.1175/1520-0450(2000)039<1185:TRIOWR>2.0.CO;2.

——, ——, ——, D. Miller, and K. E. Kunkel, 2001: A synoptic weather pattern and sounding-based climatology of freezing precipitation in the United States east of the Rocky Mountains. *J. Appl. Meteor.*, **40**, 1724–1747, https://doi.org/10.1175/1520-0450(2001)040<1724:ASWPAS>2.0.CO;2.

Razy, A., S. M. Milrad, E. H. Atallah, and J. R. Gyakum, 2011: Synoptic-scale environments conducive to orographic impacts on cold-season surface wind regimes at Montreal, Quebec. *J. Appl. Meteor. Climatol.*, **51**, 598–616, https://doi.org/10.1175/JAMC-D-11-0142.1.

Reeves, H. D., 2016: The uncertainty of precipitation-type observations and its effect on the validation of forecast precipitation type. *Wea. Forecasting*, **31**, 1961–1971, https://doi.org/10.1175/WAF-D-16-0068.1.

——, K. L. Elmore, A. Ryzhkov, T. Schuur, and J. Krause, 2014: Sources of uncertainty in precipitation-type forecasting. *Wea. Forecasting*, **29**, 936–953, https://doi.org/10.1175/WAF-D-14-00007.1.

Ressler, G. M., S. M. Milrad, E. H. Atallah, and J. R. Gyakum, 2012: Synoptic-scale analysis of freezing rain events in Montreal, Quebec, Canada. *Wea. Forecasting*, **27**, 362–378, https://doi.org/10.1175/WAF-D-11-00071.1.

Roebber, P. J., and J. R. Gyakum, 2003: Orographic influences on the mesoscale structure of the 1998 ice storm. *Mon. Wea. Rev.*, **131**, 27–50, https://doi.org/10.1175/1520-0493(2003)131<0027:OIOTMS>2.0.CO;2.

Scheuerer, M., S. Gregory, T. M. Hamill, and P. E. Shafer, 2017: Probabilistic precipitation-type forecasting based on GEFS ensemble forecasts of vertical temperature profiles. *Mon. Wea. Rev.*, **145**, 1401–1412, https://doi.org/10.1175/MWR-D-16-0321.1.

Scinocca, J. F., and Coauthors, 2016: Coordinated global and regional climate modeling. *J. Climate*, **29**, 17–35, https://doi.org/10.1175/JCLI-D-15-0161.1.

Šeparović, L., A. Alexandru, R. Laprise, A. Martynov, L. Sushama, K. Winger, K. Tete, and M. Valin, 2013: Present climate and climate change over North America as simulated by the fifth-generation Canadian Regional Climate Model. *Climate Dyn.*, **41**, 3167–3201, https://doi.org/10.1007/s00382-013-1737-5.

Smith, A., N. Lott, and R. Vose, 2011: The integrated surface database: Recent developments and partnerships. *Bull. Amer. Meteor. Soc.*, **92**, 704–708, https://doi.org/10.1175/2011BAMS3015.1.

St-Pierre, M., J. M. Thériault, and D. Paquin, 2019: Influence of the model horizontal resolution on atmospheric conditions leading to freezing rain in regional climate simulations. *Atmos.–Ocean*, **57**, 101–119, https://doi.org/10.1080/07055900.2019.1583088.

Taylor, K. E., 2001: Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res.*, **106**, 7183–7192, https://doi.org/10.1029/2000JD900719.

Uppala, S. M., and Coauthors, 2005: The ERA-40 re-analysis. *Quart. J. Roy. Meteor. Soc.*, **131**, 2961–3012, https://doi.org/10.1256/qj.04.176.

Voldoire, A., and Coauthors, 2013: The CNRM-CM5.1 global climate model: Description and basic evaluation. *Climate Dyn.*, **40**, 2091–2121, https://doi.org/10.1007/s00382-011-1259-y.