1. Introduction

Precipitation is an important meteorological climate variable and the predominant source of freshwater deposition on our planet. It is highly variable in space and time, and plays a key role in the hydrological cycle. Thus, information on precipitation is crucial for water resource management, numerical weather prediction, and transportation. Climate change has a considerable impact on precipitation and because of Arctic amplification, its ramifications are predicted to be most pronounced at higher latitudes (Bush et al. 2019; Vincent et al. 2018). Changes in precipitation have a substantial impact on the lives of Indigenous peoples, as these changes alter the composition, characteristics, and spatial and temporal occurrence of snow and ice, and introduce larger natural variability in precipitation. This challenges Inuit Qaujimajatuqangit (IQ; Inuit traditional knowledge) and the close cultural connection Inuit have with the Arctic land and wildlife. In addition, the rapidly increasing accessibility of the Canadian Arctic due to climate change makes reliable weather predictions increasingly important, particularly when it comes to navigation.

Due to sparse ground-based coverage (Mekis et al. 2018), space-borne measurements have exceptional potential in the Canadian Arctic to improve precipitation measurements and downstream products, e.g., forecasts and severe weather alerts. However, the quality of satellite-borne observations has to be assessed and possibly improved for Arctic regions before they can safely be used in further applications.

Precipitation is highly variable in space and time, which makes accurate measurements of precipitation amounts an extremely challenging task. This is even more pronounced in Arctic regions where precipitation is frequently characterized by trace amounts. Precipitation rates of less than 0.1 mm h⁻¹ are common and occur about 80% of the time (Kidd and Joe 2007), so most of the Arctic is in essence a desert. This poses particular challenges, as most instrumentation has been conceptualized for more southerly latitudes where precipitation rates up to a thousand times higher can be detected during extreme weather events. In addition, blowing snow conditions further complicate measurements. Passive microwave (PMW) sensors struggle with the detection of snow over frozen surfaces, e.g., snow and ice, which limits space-based detection abilities. These obstacles should be kept in mind when validating satellite precipitation observations with ground-based ones, as both are considerably more challenging to acquire in the Arctic, but not necessarily to a similar extent or due to the same causes.

While summer precipitation may decrease in some regions of southern Canada by the end of the twenty-first century,
increases in precipitation are generally predicted for most parts of Canada in a high greenhouse gas emission scenario (Bush et al. 2019). To confirm these model results, observations are of vital importance. However, they are not readily available for all regions. Particularly in remote areas like the Canadian Arctic, ground-based precipitation gauge measurements are very sparse. Operational weather radars are exclusively located in southern Canada (Mekis et al. 2018), so no operational radar observations are available for the Canadian Arctic. Due to this, satellite-based precipitation observations could provide essential information to fill spatial gaps in the ground-based records.

Ground-based precipitation instrumentation, data characteristics, and data management, as well as quality assessment and quality control practices employed by Environment and Climate Change Canada (ECCC) have been thoroughly documented by Devine and Mekis (2008), Mekis and Vincent (2011), and Mekis et al. (2018). While manual observations were common until the late 1990s and early 2000s, these have been gradually replaced by automated stations. Geonor T-200B and Pluvio weighing gauges have become part of the standard instrumentation for these stations since the early to mid-2000s.

The Integrated Multisatellite Retrievals for GPM (IMERG) is a multisource satellite-based precipitation retrieval algorithm (Huffman et al. 2020). It incorporates data from PMW and microwave-calibrated infrared (IR) sensors, and makes use of gauge analyses to build a comprehensive global precipitation product (Huffman et al. 2019). This study addresses the need to characterize the performance of IMERG in Arctic regions. It focuses on the final monthly product to assess the performance of the most sophisticated IMERG products compared with ground-based observations. The monthly time scale was chosen because of the low precipitation amounts in the Arctic needing a longer-term average to achieve better statistics. The data used in this study for the validation of the satellite-based IMERG observations are based on precipitation amounts recorded with ECCC’s automated Geonor and Pluvio gauges. A description of the instrumentation, data characteristics and the IMERG algorithm are provided in section 2, followed by an explanation of the methodology and statistical measures (section 3) that were applied to derive the results (section 4). The conclusions (section 5) give an overview of this study, including implications of the results and recommendations for future work.

### 2. Instrumentation, data characteristics, and the IMERG algorithm

#### a. IMERG V05 and V06

IMERG provides level 3 gridded products, which incorporates a large number of the currently available space-based PMW precipitation estimates (Huffman et al. 2018, 2019).

#### 1) CONTRIBUTING CORE AND PASSIVE MICROWAVE SATELLITES

The Global Precipitation Measurement (GPM) constellation consists of the satellites listed in Table 1, all of which occupy a low-Earth orbit (LEO). It should be noted that the Tropical Rainfall Measuring Mission (TRMM) Microwave Imager (TMI) (Kummerow et al. 1998) was not included in V05 but is part of the PMW ensemble utilized for V06. Furthermore, contributions from geosynchronous-Earth-orbit (GEO) IR satellites are included to provide additional information and to fill gaps between LEO satellite paths where information from other sources is not sufficient. IMERG is calibrated to the combined Dual-Frequency Precipitation Radar (DPR)/GPM Microwave Imager (GMI) product for the GPM era and the combined product of TRMM for its respective era.

The Global Precipitation Measurement Core Observatory (GPM-CO) has been collecting measurements from 12 March 2014 to present and covers a latitudinal range from 25°S to 25°N, compared to 35°S–35°N for TRMM. Its payload consists of the DPR (Hamada and Takayabu 2016) and the GMI (Draper et al. 2015). The nadir-viewing DPR is composed of Ku-band (13.6 GHz) and Ka-band (35.5 GHz) channels (Iguchi et al. 2018) with swaths of 245 and 120 km respectively. The additional Ka-band channel of DPR provides significant improvements over its predecessor TRMM Precipitation Radar (PR) (13.8 GHz) (Kummerow et al. 1998): It gives information on the drop size distribution (DSD), which is crucial for the conversion of reflectivity to rain fall rate, and varies with region, season, and rain type (e.g., solid versus liquid precipitation). In addition, it enables more accurate estimates of the altitude where the precipitation phase changes (bright band) and thus allows for better corrections. The GMI is a microwave radiometer...
that features 13 channels ranging from 10 to 183 GHz. Thus, it is capable of detecting light, moderate, and heavy precipitation, with high-frequency channels (166 and 183 GHz) specifically targeting light precipitation, small ice particles, and snowfall over snow-covered land (Hou et al. 2014). It scans a 904-km-wide swath in a conical fashion at an off-nadir angle of 48.5°, which leads to an approximate time lag of 67 s with the DPR.

PMW sensors such as GMI, TMI, the Advanced Technology Microwave Imager/Sounder (ATMS), the Special Sensor Microwave Imager/Sounder (SSMIS), and the Advanced Microwave Scanning Radiometer 2 (AMSR-2) employ imager channels (i.e., 10–89 GHz), which are considered most suitable for the moderate and heavy precipitation associated with low and midlatitudes. High-frequency channels (i.e., above 89 GHz) as deployed on GMI, ATMS, and the Microwave Humidity Sounder (MHS) allow for the sensing of light precipitation as well as snow and ice particles that are present at higher latitudes.

2) ALGORITHM AND DATA CHARACTERISTICS

IMERG provides global precipitation estimates at a spatial resolution of 0.1° × 0.1°. Products from an early, a late, and a final run are available with latencies of 4 h, 14 h, and 2.5 months, respectively. For the final run, both half-hourly and monthly products are available, while early and late data are provided solely on a half-hourly basis. The final run provides the most advanced estimate and incorporates monthly gauge data provided by the Global Precipitation Climatology Centre (GPCC) [Rudolf et al. (2011) and Schneider et al. (2018)] for V05 and V06, respectively for bias correction to create research-level precipitation estimates.

The Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks with Cloud Classification System (PERSIANN-CCS) algorithm (Hong et al. 2004) is used to dynamically intercalibrate IR satellites estimates with the IMERG PMW estimates for latitudes between 60°S and 60°N. Merged IR brightness temperatures at 4 km are used to calculate motion vectors for the Climate Prediction Center (CPC) morphing–Kalman filter (CMORPH-KF) algorithm (Joyce and Xie 2011), where PMW precipitation rate estimates closest to the adjacent analysis time are propagated both forward and backward from their observation time toward that time using cloud system advection vector (CSAVs) derived from GEO-IR images (for IMERG V05). By inversely weighting the forward and backward PMW estimates with their error variance, a forecast of the precipitation analysis can be defined. In cases where the gap between two PMW measurements exceeds 90 min, IR measurements are used as a fallback option since they have higher correlation than the propagated PMW estimates. Following the same approach as used in the algorithm for GPM's predecessor TRMM (Huffman et al. 2007), the satellite precipitation estimates are bias corrected with monthly GPCC gauge data. Prior to use, the gauge data are corrected for wind undercatch by multiplication with climatological adjustment ratios from Legates and Willmott (1990) as part of the IMERG processing. A multisatellite estimate is then adjusted to the large-spatial-scale mean of the gauges and subsequently an error-weighted estimate of the gauge and satellite estimates is calculated.

The full IMERG V05 algorithm can be applied for latitudes between 60°S and 60°N. However, regions like the Canadian Arctic (i.e., latitudes north of 60°N) are not covered by geostationary IR satellites and thus make the application of CMORPH-KF impossible. Hence, only merging is applied for these latitudes in IMERG V05. The aforementioned obstacle has been overcome in V06 by using the total precipitable water vapor (TQV) field from the global atmospheric reanalysis Modern Era Retrospective Reanalysis 2 (MERRA-2) (not near real time) and the global forecast analysis Goddard Earth Observing System model Forward Processing (GEOS-FP) (near real time) to calculate displacement vectors in IMERG V06 (Tan et al. 2019). Due to this improvement, morphing is applied in Arctic regions in IMERG V06 but is confined to non-snowy or non-icy surfaces. The National Oceanic and Atmospheric Administration (NOAA) Autosnow product (Romanov 2016) was used to mask respective areas.

This study focuses exclusively on the final monthly data products of IMERG V05 and V06 and is confined to the Canadian Arctic (latitudes north of 60°N).

b. Independent ECCC climate archive stations

For this study, stations have been selected from the ECCC Climate Network that are not ingested into the GPCC data product (Fig. 1; stations that temporarily overlap only with IMERG V06 are indicated by red crosses). The latter is of particular importance since GPCC data are used to remove biases in the final IMERG products. Thus, by only selecting stations that are not ingested into GPCC, it is ensured that the comparison is performed with an independent dataset. Table 3 provides an overview of the selected stations, their three-letter Transport Canada Identifier (TC IDs) as well as their associated latitudes and longitudes. Stations that temporally only overlap with IMERG V06 are indicated in bold.
1) INSTRUMENTATION

Geonor and Pluvio weighing gauges are used to automatically collect precipitation data at the surface (Mekis et al. 2018). The Geonor T-200B and the Pluvio1 employed have a capacity of 600 and 1000 mm, respectively (Devine and Mekis 2008), and were introduced for operational use in Canada between 2002 and 2006. Recently, the single-alter-shielded Geonor gauges have been replaced with double-alter-shielded Pluvio2 gauges throughout the network. The dry bulb temperatures provided at the same locations are used to distinguish between precipitation types. These temperatures are solely used for filtering outliers from the hourly gauge data, which is described in detail in section 3b.

2) DATA CHARACTERISTICS

For this study, hourly ECCC Climate Network station precipitation data are used, which are available under element number 262 in the ECCC climate archive (Environment and Climate Change Canada 2018). Initially, no quality checks were performed at the ingest stage (indicated by "R" for "raw"), while basic automatic quality assessment (indicated by "Q") was introduced at the ingest stage on 10 December 2013. Precipitation amounts are provided in millimeters with a step size of 0.1 mm and reported for every hour, i.e., for minutes 0–60. ECCC Climate Network hourly dry bulb temperatures are available in the ECCC climate archive under element number 078 and feature a step size of 0.1°C.

3. Methodology

a. Distinction between liquid, mixed, and solid precipitation

Since the ground-based measurements of liquid, mixed, and solid precipitation are evidently impacted differently by factors such as wind, a distinction is made between these three types. To achieve this, hourly temperature data from each ground-based station are used following the methodology of Kochendorfer et al. (2017) and Wolff et al. (2015): when dry bulb ground temperatures fall below −2°C precipitation is assumed to be mainly solid. For temperatures above +2°C, mainly liquid precipitation is expected, while temperatures between −2° and +2°C lead to mixed precipitation containing both solid and liquid hydrometeors. After removing probable outliers from the ECCC Climate Network hourly precipitation data (see section 3b), monthly mean precipitation rates (mm h⁻¹) are calculated and subsequently compared to IMERG final monthly precipitation rates.

b. Outlier removal in the ground-based data

The ECCC Climate Network station hourly precipitation data occasionally exhibit unphysical outliers, i.e., easily exceeding 100 mm h⁻¹, while hourly precipitation rates in the Arctic are usually close to 0.1 mm h⁻¹. To resolve this issue,
TABLE 3. ECCC Climate Network stations used in this study. Stations that only overlap with IMERG V06 are indicated in bold.

<table>
<thead>
<tr>
<th>TC ID</th>
<th>Station name</th>
<th>Latitude</th>
<th>Longitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>YPX</td>
<td>Puvirnituq A</td>
<td>60.05</td>
<td>-77.29</td>
</tr>
<tr>
<td>YEK</td>
<td>Arviat A</td>
<td>61.09</td>
<td>-94.07</td>
</tr>
<tr>
<td>XLL</td>
<td>Lindburg Landing</td>
<td>61.12</td>
<td>-122.85</td>
</tr>
<tr>
<td>XDK</td>
<td>Deadman Valley</td>
<td>61.26</td>
<td>-124.47</td>
</tr>
<tr>
<td>VBR</td>
<td>Burwash Airport</td>
<td>61.37</td>
<td>-139.02</td>
</tr>
<tr>
<td>YDB</td>
<td>Burwash A</td>
<td>61.37</td>
<td>-139.04</td>
</tr>
<tr>
<td>WBJ</td>
<td>Inner Whalebacks</td>
<td>61.92</td>
<td>-113.73</td>
</tr>
<tr>
<td>MFX</td>
<td>Salluit</td>
<td>62.18</td>
<td>-75.67</td>
</tr>
<tr>
<td>WMT</td>
<td>Lac La Martre</td>
<td>63.13</td>
<td>-117.24</td>
</tr>
<tr>
<td>NJS</td>
<td>Coral Harbour RCS</td>
<td>64.18</td>
<td>-83.35</td>
</tr>
<tr>
<td>YWE</td>
<td>Wekweeti A</td>
<td>64.19</td>
<td>-114.08</td>
</tr>
<tr>
<td>ZGH</td>
<td>Fort Good Hope CS</td>
<td>66.24</td>
<td>-128.64</td>
</tr>
<tr>
<td>YVL</td>
<td>Colville Lake A</td>
<td>67.02</td>
<td>-126.13</td>
</tr>
<tr>
<td>YVM</td>
<td>Qikiqtarjuaq A</td>
<td>67.55</td>
<td>-64.03</td>
</tr>
<tr>
<td>YHK</td>
<td>Gjoa Haven A</td>
<td>68.64</td>
<td>-95.85</td>
</tr>
<tr>
<td>XTV</td>
<td>Trail Valley</td>
<td>68.75</td>
<td>-133.50</td>
</tr>
<tr>
<td>YUX</td>
<td>Hall Beach A</td>
<td>68.78</td>
<td>-81.24</td>
</tr>
<tr>
<td>WOI</td>
<td>Ivavik Nat. Park</td>
<td>69.16</td>
<td>-140.15</td>
</tr>
<tr>
<td>YCY</td>
<td>Clyde River A</td>
<td>70.49</td>
<td>-68.52</td>
</tr>
<tr>
<td>NFR</td>
<td>Fort Ross</td>
<td>72.02</td>
<td>-94.20</td>
</tr>
<tr>
<td>YOI</td>
<td>Pond Inlet A</td>
<td>72.69</td>
<td>-77.97</td>
</tr>
<tr>
<td>WYF</td>
<td>Thomsen River</td>
<td>73.23</td>
<td>-119.54</td>
</tr>
<tr>
<td>NPL</td>
<td>Cape Liverpool</td>
<td>73.67</td>
<td>-78.29</td>
</tr>
<tr>
<td>NPV</td>
<td>Cape Providence</td>
<td>74.46</td>
<td>-112.16</td>
</tr>
</tbody>
</table>

different approaches are possible: simply applying an upper threshold would remove drastic outliers, however, it is also desirable to eliminate probable outliers (Tukey 1977) from the dataset, while still retaining a maximum number of measurements. Tukey’s boxplot method (Tukey 1977) provides a good measure for outlier detection for normally distributed samples, but it reaches its limits when it comes to the skewed distributions that are typical of precipitation measurements. In these cases, the adjusted boxplot suggested by Hubert and Vandervieren (2008) is better suited to identify outliers and conserves more data points by taking the skewness of the distribution into account, i.e., the long tail of hourly precipitation amount distributions. The adjusted boxplot utilizes the medcouple (MC) index (Brys et al. 2004) as a robust measure for the skewness of a distribution which is defined as follows:

\[ MC = \text{med}_{x_i \neq x_j} h(x_i, x_j), \]  

with the sample median \( \text{med}_{x_i \neq x_j} \) and for all \( x_i \neq x_j \), the kernel function \( h \) is given by

\[ h(x_i, x_j) = \frac{(x_j - Q_1)(x_i - Q_3)}{x_j - x_i}. \]  

The medcouple always lies between \(-1\) and \(1\), with skewed left distributions (e.g., precipitation) indicated by negative values and skewed right distributions indicated by positive values. In the case of a symmetric distribution, the medcouple is zero.

While the medcouple is a robust measure for the skewness of a distribution it encounters difficulty when the sample’s distribution is poorly resolved. In this study, the coarse resolution of the ground-based instrumentation with respect to the quantity of interest, i.e., hourly precipitation amounts in Arctic regions, results in poorly defined distributions. This, in turn, results in medcouple estimates that are not representative and, ultimately, in unfavorable cutoff values. To overcome this problem, a gamma distribution (Thom 1958) is iteratively fitted to the data. Subsequently, a finer-resolved random dataset is created that follows the previously determined gamma distribution. The medcouple is then calculated from that artificially created distribution. This procedure is performed for each station individually and for each precipitation type at that station. Cutoff levels \((x_{cut-off})\) for probable outliers (for a skewed left distribution) are calculated as follows (Brys et al. 2004):

\[ x_{cut-off} = \begin{cases} 
Q_1 - 1.5e^{-3MCQR} & \text{lower cut-off} \\
Q_3 + 1.5e^{3MCQR} & \text{upper cut-off}
\end{cases} \]  

with the first and third quartiles \( Q_1 \) and \( Q_3 \), respectively, and the interquartile range \( \text{IQR} = Q_3 - Q_1 \). The upper cutoffs for solid, mixed, and liquid precipitation are then applied to the original dataset of each station. In cases where the mixed or solid cutoff is higher than the liquid cutoff, the liquid cutoff is applied instead.

c. Comparison of time series and statistical measures

To give an overview of how the IMERG final monthly products compare with ground-based Climate Network station measurements, the mean bias \( B_M \) [Eq. (4)] is calculated for each station location as follows:

\[ B_M = \frac{1}{N} \sum_{i=1}^{N} \frac{x_{IMERG, i} - x_{CNS, i}}{N} \]  

with the individual month \( i \), the total number of months \( N \) included in the comparison, and the respective monthly precipitation rates for IMERG \( x_{IMERG, i} \) and the ECCC Climate Network stations \( x_{CNS, i} \). In addition, the mean relative bias has been calculated follows:

\[ B_R = \frac{1}{N} \sum_{i=1}^{N} \frac{x_{IMERG, i} - x_{CNS, i}}{\frac{x_{IMERG, i} + x_{CNS, i}}{2}} \times 100. \]  

Since temporal coverage for some of the ground-based stations was found to be highly variable and to account for the spatial coverage of IMERG when comparing areas, we also calculate the difference weighted by the error estimates as follows:

\[ B_W = \epsilon_w \sum_{i=1}^{N} \frac{x_{IMERG, i} - x_{CNS, i}}{\sigma^2_{IMERG, i} + \sigma^2_{CNS, i}} \]  

with the number of months \( N \), the standard errors of the mean.
\[
\sigma_{x_{\text{IMERG}}} = \frac{\sigma_{\text{IMERG}}}{\sqrt{N}} \\
\sigma_{x_{\text{CNS}}} = \frac{\sigma_{\text{CNS}}}{\sqrt{N}}
\]  

for IMERG and the Climate Network stations, respectively, and the weighting coefficient

\[
\omega = \frac{1}{\sum_{i=1}^{N} \frac{1}{\sigma_{x_{\text{IMERG}}}^2 + \sigma_{x_{\text{CNS}}}^2}}.
\]

The standard error of the mean is used for the ground-based stations and for IMERG estimates based on areas. For comparisons with the closest IMERG grid point, we reverted to the error estimate given by the product. In the very few cases where the standard error was zero, due to a small sample size and all of the collected precipitation estimates being zero, a new standard error was calculated using the mean standard deviation for the respective station [Eq. (9)] rather than using the standard deviation for the individual month

\[
\sigma_{M}^2 = \frac{1}{N} \sum_{i=1}^{N} \sigma_{x_{i}}^2.
\]

This approach is used to avoid giving particularly high weighting to these data points, since the remaining weighting would be solely based on the other instrument’s standard error of the mean and thus the combined error would be considerably smaller. Using the mean standard deviation in these special cases provides the most neutral approach.

Furthermore, the Pearson correlation coefficient \( r \) was calculated for each station:

\[
r = \frac{\sum (x_{\text{IMERG}} - \bar{x}_{\text{IMERG}})(x_{\text{CNS}} - \bar{x}_{\text{CNS}})}{\sqrt{\sum (x_{\text{IMERG}} - \bar{x}_{\text{IMERG}})^2 \sum (x_{\text{CNS}} - \bar{x}_{\text{CNS}})^2}}.
\]

A breakdown into seasons is also performed to examine the agreement between the IMERG and ground-based measurements during mainly snow-free and snow-covered times of year.

4. Results

4.1. Mean bias and error weighting

Monthly precipitation rates for the ground-based measurements were calculated from hourly rates. For IMERG the mean monthly precipitation rates over areas enclosing the ground-based stations were calculated, while gridded monthly rates for IMERG are readily available. For the months of November through April, the IMERG precipitation estimates consist mainly of gauge analyses for both data versions, which corresponds to approximately 50% of the data north of 60°N. In case of the ground-based stations, irregular sampling in time can lead to biases in the monthly means. For IMERG irregular sampling within the respective areas can also induce a bias. To illustrate the impact of irregular sampling based on the aforementioned averaging, an overview of the mean bias and the mean bias based on error-weighted monthly means between IMERG precipitation rate and that for each of the ground-based stations is shown in Figs. 2 and 4, respectively. The bias is given as a mean monthly precipitation rate (mm h\(^{-1}\)). The relative bias is shown in Fig. 3. The comparisons are sorted by latitude ranging from Puvirnituq (YPX) at 60.05°N at the bottom to Cape Providence (NPV) at 74.46°N at the top of each bar graph. Each triplet of bars (blue, red, and orange) represents the mean bias (Fig. 2) or mean error-weighted bias (Fig. 4) at the respective station. The blue bar refers to the comparison with the grid point in IMERG that is closest to the respective ground-based station. The red and the orange bars are based on comparisons with means of the IMERG data that were derived from an area around each ground-based station. The red bar is associated with a grid box defined by ±1° latitude and ±2° longitude centered around the ground-based station. The orange bar represents a comparison based on a box with ±110 km around the ground-based station.

Since different versions of GPCC were used as well as different snow cover products, and potentially due to changes in the IMERG algorithm itself, gaps in the IMERG data appear in different locations for V05 and V06. Due to this, a different number of months might contribute to the comparison for each location between IMERG V05 and V06. Hence, the number of months contributing to the estimate of the mean bias and the mean error-weighted bias is indicated by the number to the right of each bar. In Figs. 2 and 4 the panels (a), (b), and (c) show comparisons of the ground-based measurements with IMERG V05, V06 for the same period as V05 (12 March 2014 to 30 June 2018) and the full V06 time period (12 March 2014 to 31 March 2019), respectively. Even though the same time period is considered in panels (a) and (b), the number of months contributing to the estimated bias might still differ between IMERG V05 and V06 due to different spatial coverage of the two versions as mentioned above (cf. Fig. 5). Figure 5 also displays how over some stations IMERG only provides data during some months, e.g., over Cape Providence (NPV) data are only available during the summer months likely due to the surface being covered in ice and snow during other times of the year.

In general, the mean bias between IMERG and the ground-based precipitation rate is very similar for the two versions of IMERG at each location. Exceptions are Colville Lake (YVL) closest (blue bar), and Oigikotjarjaq (YVM) and Clyde River (YCY) for the two areas (orange and red bar) around the station. This is considerably more pronounced in the mean bias (Fig. 2) than in the error-weighted bias (Fig. 4), with there being almost no difference in the mean error weighted bias based on IMERG areas. For Colville Lake (YVL) and the closest grid point in IMERG, the difference between IMERG V05 and V06 decreases noticeably when error weighting is applied. This indicates that most of the difference in the comparisons of IMERG with these three ground-based stations was induced by a sampling bias in either the ground-based measurements or IMERG. However,
it is impossible to tell whether this results primarily from a temporal sampling bias in the ground-based data or from a spatial sampling bias in IMERG, or whether there is a significant contribution from both sources. This might also change from station to station. For YCY, the comparison to the closest IMERG grid point also shows a noticeable difference in the mean bias, however, a different number of months goes into this comparison due to different spatial gaps in IMERG.
V05 and V06 (50 versus 44 for V05 and V06 during the same period as V05). Other than providing comparisons with two additional stations, the comparison with the full time series of IMERG V06 shows small differences compared to the shorter time series of the same version, which suggests that the assessed biases are overall rather robust.

For most stations, IMERG reports higher precipitation rates than the ground-based stations, a feature that is even more consistent in the mean error-weighted bias. This result is consistent with the finding that precipitation weighing gauges, such as Geonor and Pluvio, may underestimate precipitation considerably due to wind-associated undercatch.

**FIG. 3.** Mean relative bias in precipitation rate of IMERG final monthly compared to ECCC Climate Network stations: (a) IMERG V05, (b) IMERG V06 during the same time period as V05 (12 Mar 2014–June 2018), and (c) IMERG V06 for the full V06 time period. The numbers next to each bar represent the number of months contributing to the respective bias.
evaporation and wetting losses (Mekis and Vincent 2011). Underestimation of precipitation measurements due to these effects can easily exceed 20% in the Canadian Arctic. Respective adjustments have not been applied to the ground-based data considered in this study as they are still being developed.

Overall, the derived bias indicates that the differences between the IMERG precipitation rates and the ground-based ones are larger at lower Arctic latitudes. This can be attributed to the fact that precipitation rates generally decrease with increasing latitude, so that the absolute differences are smaller at higher latitudes. Figure 3 supports this since a decrease with increasing latitude is not obvious in the relative bias. Mean biases between IMERG and the ECCC Climate Network stations (Fig. 2) range up to 0.05 mm h\(^{-1}\).
While error-weighted biases (Fig. 4) range up to 0.04 mm h$^{-1}$ (equivalent to 0.96 mm day$^{-1}$). With increasing latitude, a difference in the bias for the two areas ($\pm1^\circ$ latitude/$\pm2^\circ$ longitude; $\pm110$-km box) becomes slightly more noticeable. This is to be expected since the area selected based on latitude and longitude becomes increasingly distorted at higher latitudes compared to the area based on a fixed distance from the ground-based station. Nevertheless, for most stations, both area methods yield very similar results while the biases based on the closest IMERG grid point can considerably deviate from the area-based biases. The latter is less pronounced in IMERG V06 for most stations north of 65°N, particularly YVL, YVM, and YCY in Fig. 2. This might originate in the CMORPH-KF algorithm being applied north of 60°N in IMERG V06. For future investigations, it seems appropriate, due to the similarity of the two methods of area selection, to only consider the less computationally demanding approach for analyses, i.e., selecting based on the IMERG grid.

**b. Spatial dependencies**

Figure 6 shows the error-weighted bias of the comparison with the closest IMERG point in IMERG V05 and IMERG V06 (entire period) in the top and bottom panel, respectively. A similar comparison is depicted in Fig. 7 for the latitude–longitude-based area. It should be noted that the point in southwestern Yukon is actually an overlap of two stations (VBR and YDB) located almost entirely in the same location. No clear pattern based on coastal versus continental stations can be seen in these maps. This holds for both versions, V05 and V06 of IMERG and also both for the closest IMERG point as well for the IMERG data in an area around the ground-based stations.

**c. Correlations and seasonal dependencies**

Figure 8 provides further insight into the comparisons of the closest grid point in IMERG with two of the ground-based stations, Deadmen Valley (XDK) and Wekweeti A.
These two stations were chosen to demonstrate the advantage of the error-weighted approach. Furthermore, the color coding provides additional means of interpretation, since it indicates the probability of liquid precipitation (0 = entirely solid; 1 = entirely liquid; Huffman et al. 2018, 2019). Error bars in the horizontal direction represent the standard error of the mean for the ground-based measurements, while error bars in the vertical direction represent the errors associated with the monthly estimated precipitation rate of the closest IMERG grid point. Deadmen Valley shows a case where there is a stark outlier in the ground-based dataset, while for Wekweeti the outlier occurs in the IMERG dataset. The outlier in the ground-based data is sampling related, i.e., the number of data points is rather small and the distribution of these points is not representative for this station. This kind of outlier must not be confused with the outliers of the original hourly data that were filtered out in a previous step as discussed in section 3b. That outlier removal addresses unphysical values in the hourly data, while the outliers in the monthly means are simply unrepresentative due to the aforementioned sampling bias. The outlier in IMERG data is more difficult to attribute to a specific issue. However, when investigating these plots for all available locations, outliers falling into the solid precipitation regime were found to embody the vast majority of anomalies. This confirms that measurements are more challenging during the period of the year where...
solid precipitation occurs for both IMERG and ground-based precipitation gauges. In both cases in Fig. 8, the error bars are large for the outliers. Hence, outliers are weighted less strongly than the other data points, providing a more realistic bias estimate without throwing out data points. Since precipitation rates tend to be slightly smaller during colder seasons, the associated deviations could theoretically also induce changes when calculating the mean error-weighted bias. However, this effect was found to have a minor impact and can be considered negligible.

A seasonal breakdown of the error-weighted bias can be found in Fig. 9. Both versions of IMERG show smaller biases with respect to the ground-based precipitation rate during winter (DJF) and spring (MAM). These are the colder Arctic seasons during which less precipitation is expected in general. Accordingly, differences can also be expected to be smaller during these times of the year. Similar explanations apply for the wider variability in summer (JJA) and fall (SON) seasons, with there being larger amounts of precipitation expected in general but also a larger variability which is reflected in the biases.

Figure 10 shows the seasonal correlation coefficients for IMERG versus the ECCC Climate Network monthly precipitation rates. The left and right panels show the results for the different versions of IMERG, V05 and V06 for its full period, respectively. Both versions of IMERG have stronger
correlations with the ground-based measurements during the summer (JJA) and fall (SON). This is potentially caused by IMERG estimates comprising of gauge measurement to a large extent during these months. IMERG V05 shows higher variability in correlation coefficients than IMERG V06. This might be attributed to CMORPH-KF being applied in regions north of 60°N in IMERG V06. The strongly negative correlation coefficient during the spring season (MAM) in the comparison of IMERG V06 with the ECCC Climate Network stations originates from there only being very few months (three data points) available during that season for the ±110-km area around Cape Providence (NPV). The distribution of these few points leads to a strongly negative correlation (−0.99). In the comparison with IMERG V05 this does not stand out, because the three data points are distributed slightly differently, leading to correlation coefficient of −0.53. Neither of these correlation coefficients for Cape Providence (NPV) should be considered representative, since they originate from a very small sample size.

5. Conclusions

In this study, versions V05 and V06 of IMERG final monthly product are compared with 25 ECCC Climate Network stations in the Canadian Arctic, i.e., north of 60°N, to assess the performance of the satellite-based level 3 products. Particular emphasis is placed on the selection of the ground-based stations to ensure that they are entirely independent from the IMERG products.

All comparisons were performed using three approaches to select IMERG data: the first uses the IMERG data at the grid point that is spatially closest to the respective station, while
the other two methods use the mean IMERG precipitation rate estimated from two areas around each station. One of these areas uses the IMERG latitude–longitude grid for the selection of a grid box, which is 2° wide in latitude and 4° wide in longitude and centered on the respective ECCC Climate Network station. The second area is based on a square with a side length of 220 km centered on the respective ground-based station. These comparisons are associated with the same colors throughout this document, i.e., blue for the comparison with the grid point closest to the ground-based station, red for the selection based on the native IMERG latitude–longitude grid, and orange for the selection based on a 220-km square around the station.

The mean bias and mean error-weighted bias show that, at most locations, IMERG observes higher precipitation rates than the ground-based ECCC Climate Network stations. Mean biases of up to 0.05 mm h⁻¹ and mean error-weighted biases of up to 0.04 mm h⁻¹ were found in this study. The detection of more precipitation by IMERG compared to the ECCC Climate Network ground-based stations is consistent with the understanding that precipitation weighing gauges tend to underestimate precipitation amounts due to wind-associated undercatch, evaporation and wetting losses (Mekis and Vincent 2011). This emphasizes the importance of efforts currently underway to correct for these issues, e.g., the development and application of transfer functions based on the World Meteorological Organization (WMO) Canadian Solid Precipitation Experiment (C-SPICE) campaign (Pierre et al. 2019).

With few exceptions, the mean bias and mean error-weighted bias of IMERG relative to the Climate Network precipitation rates show similar results for both versions of IMERG, V05 and V06. In addition, the results are similar for comparisons of V06 for the time period covered by V05 (March 2014–June 2018) and the full V06 time series (March 2014–February 2019), which indicates that the assessed findings for V06 are robust.

Looking at the comparisons with different means of selecting IMERG data, those using the closest point in IMERG sometimes differ considerably from those using areas means. This is likely related to the highly variable nature of precipitation, even on a monthly scale. For the comparisons based on the two different areas, the assessed biases were similar in most cases. Thus, for future work, comparisons solely based on a latitude–longitude area selection might be sufficient and more desirable, since the selection based on a geospatial distance is more computationally demanding.

Noticeable differences between the mean bias and the mean error-weighted bias occur mainly due to sampling irregularities for the ground-based stations and could not be attributed to a specific challenge that IMERG or contributing data sources face. However, the vast majority of outliers for both IMERG and ground-based precipitation rates fall into the solid precipitation regime, which likely points to the difficulty
of obtaining measurements during colder seasons, e.g., when the precipitation rates are low and there is snow or ice on the ground.

While the ground-based stations are rather evenly spread out over the Canadian Arctic, no dependency of the bias could be found for coastal stations versus continental ones. This holds for both IMERG V05 and V06.

To gain a better understanding of IMERG’s performance based on season, the mean error-weighted bias and the correlation coefficient were assessed separately for winter (DJF), spring (MAM), summer (JJA), and fall (SON). Biases between IMERG and ground-based precipitation rates are smaller during winter and spring than during summer and fall. Since precipitation amounts tend to be higher during summer and fall, similar patterns are to be expected for biases. Both versions of IMERG, V05 and V06, have stronger correlations with the ground-based data in summer and fall, with correlation coefficient medians around 0.75–0.8. Both versions of IMERG exhibit correlation coefficient medians below 0.5 for winter and spring. This may be related to the issues that PMW sensors have over icy and snowy surfaces and to the significantly increased problems that weighing gauges have with undercatch once there is solid precipitation instead of liquid. Over icy and snowy surfaces PMW sensors are not used in the algorithm leading to IMERG values strongly dominated by ground-based sources. These are, however, sparse and IMERG estimates are correspondingly not very accurate in locations that are far away from ground-based stations used in the algorithm. IMERG V06 shows slightly lower median correlation coefficients with the ground-based measurements for winter and spring.

This study provides a first assessment of IMERG precipitation rates in the Canadian Arctic. The use of IMERG data to improve certain applications, particularly near-real-time ones, evaluation of the other IMERG products would be beneficial. Comparison of both the final and early run half-hourly IMERG products with the ECCC Climate Network stations and instruments at ECCC supersites like Iqaluit (Mariani et al. 2016; Joe et al. 2020) could provide a more in-depth understanding of the IMERG dataset in the Arctic. Comparisons with ECCC research radars in Iqaluit and Whitehorse would be of particular interest, since they provide spatial resolution that comparisons with point sources like weighing gauges cannot provide.

Acknowledgments. This work has been supported by the University of Toronto’s Arts and Science Postdoctoral Fellowship Program, the University of Toronto, and the Natural Sciences and Engineering Research Council of Canada (RGPIN-2017-06664). The authors thank the Meteorological Service of Canada (MSC) for providing precipitation data, and MSC operators for collecting...
the precipitation data. Furthermore, the authors thank Thomas von Clarmann (Karlsruhe Institute of Technology) and Dominique Brunet (Environment and Climate Change Canada) for their insights on statistics and fruitful discussions.

Data availability statement. The Integrated Multisatellite Retrievals for GPM (IMERG) data used in this study are available under http://dx.doi.org/10.5067/GPM/IMERG/3B-MONTH/05. These measurements of Environment and Climate Change Canada (ECCC) Climate Network stations are available from ECCC’s National Climate Data and Information Archive Historical Climate Data portal at various temporal scales ranging from hourly to annual (https://climate.weather.gc.ca/, ECCC 2018).

REFERENCES


