Monitoring the impacts of weather radar data quality control for quantitative application at the continental scale

Daniel Michelson1 | Bjarne Hansen2 | Dominik Jacques2 | François Lemay2 | Peter Rodriguez3

1Environment and Climate Change Canada, Toronto, Ontario, Canada
2Environment and Climate Change Canada, Dorval, Québec, Canada
3Environment and Climate Change Canada, King City, Ontario, Canada

Correspondence
Daniel Michelson, Environment and Climate Change Canada, 4905 Dufferin Street, Toronto, ON M3H 5T4, Canada.
Email: daniel.michelson@canada.ca

Abstract
As part of a suite of quality control methods applied to Canadian and American weather radar data before their assimilation into a numerical weather prediction model, the combination of thresholded depolarization ratio and a speckle filter was applied to American data with the purpose of identifying and removing non-precipitation echoes. This polarimetric quality control replaces a set of image-analysis-based methods used in a previous study and based on reflectivity information only. The old and new quality-controlled results were objectively assessed using meteorological aerodrome report (METAR)-based precipitation occurrence observations and a set of five common contingency table skill scores with all available Next Generation Weather Radar (NEXRAD) Level II data from the contiguous United States for August 2016. The new quality control yields consistently improved skill scores, indicating higher quality radar data for downstream application. The process whereby the radar data are quality controlled and assessed comprises a framework with the ability to monitor the impacts of quality control to radar data quality over time. In turn, this allows for the introduction of changes to data acquisition and processing with the ability to monitor the impacts on data quality: a scientific evidence-based quality assurance process as part of change management.

KEYWORDS
data quality, meteorological aerodrome report (METAR), monitoring, quality assurance, quality control, weather radar

1 | INTRODUCTION

The scientific field of radar meteorology has struggled since its inception with the challenge of identifying and separating meteorologically relevant echoes from those not relevant to a given meteorological application area. It is ironic that the field itself is a by-product of the military use of radar during the Second World War (e.g. Fabry, 2015), where precipitation was considered a source of contamination.

The quality control of weather radar data has been a major topic within the field. Many methods have been developed over the years to identify and suppress clutter from land and sea (Riedl, 1995), anomalous propagation echoes (Alberoni et al., 2001), biological targets (mostly birds and insects) (Zrnić and Ryzhkov, 1999), radio-
frequency interference (Saltikoff et al., 2016), and second-trip echoes (Park et al., 2016), to correct for non-meteo-
logical artefacts such as beam blockage due to topogra-
phy (Bech et al., 2007), and to correct for meteorological
artefacts such as attenuation of the radar signal by pre-
cipitation (Hitschfeld and Bordan, 1954). Systematic
quality control chains, including identification and the
removal of non-precipitation echoes, were implemented
in national weather radar networks in the Nether-
lands (Wessels and Beekhuis, 1994), United States (Fulton et al., 1998), UK (Harrison et al., 2000), Switzerland
(Germann and Joss, 2004), Poland (Ośródka et al., 2012),
and Japan (JMA, 2017), among other places.

As part of European Co-operation in the Field of Sci-
cientific and Technical Research (COST) Action 717 (Rossa et al., 2005), entitled “Use of Radar Observations in
Hydrological and NWP Models”, weather radar quality
control and characterization were addressed systemati-
cally and comprehensively across Europe (Michelson et al., 2005). The challenge of identifying and removing
non-precipitation echoes, as a critical step in systematic
production chains, is addressed for each country across
Europe, including references to documentation in which
the methods used are published. This led to recommen-
dations on how to represent data quality information
with operational weather radar data exchanged interna-
tionally (Hollemann et al., 2006), implementation in data-
processing software (e.g. Michelson et al., 2018), and inte-
gration into the production of continental-scale European
weather radar products (Huuskonen et al., 2014).

With the ongoing global rollout of new operational
weather radars with polarimetric capabilities, the current
focus is on benefiting from the potential improvements
to data quality by exploiting the information content in
the polarimetric data. Radar target identification has
been developed for simpler separation of meteorological
from non-meteorological echoes (e.g. Gourley et al., 2007;
Ye et al., 2015; Kilambi et al., 2018) purely for the sake of
quality control. More sophisticated hydrometeor classi-
fication approaches (e.g. Vivekanandan et al., 1999; Park et al., 2009) can include several non-meteorological classes
useful for this form of quality control, while the main
purpose is to locate different precipitation types, for
example hail for warnings. Recent advances in radar
ornithology have also contributed towards improved
quality control (Lin et al., 2019) for the separation of
weather and biological echoes.

The context of the present study is that of the assimili-
ation of weather radar data into numerical weather pre-
diction (NWP) systems. Introducing continental-scale
radar data assimilation to Environment and Climate
Change Canada (ECCC) (Jacques et al., 2018) gave a
starting point for such activities with improved forecast

skill, and part of this starting point was the weather radar
data quality control-processing chain using all the opera-
tional weather radar data available from the Canadian
and American national networks. At the time the first
impact studies were conducted, it was felt that the quality
controls were applied aggressively, removing non-preci-
ipitation echoes at the expense of a noticeable amount of
real precipitation. However, due to of the nature of the
latent heat nudging (LHN) method used with the assimil-
ated precipitation inferred from radar reflectivities, this
loss of precipitation was considered acceptable due to of
LHN’s sensitivity to non-precipitation (false-positive)
information that risked being left in the radar data had
the methods been used less aggressively.

Our intent has been to replace the aggressive removal
of precipitation by an algorithm that could better discrim-
inate precipitating from non-precipitating echoes. In this
study, such a weather radar quality control algorithm is
applied, based on the derived depolarization ratio (DR)
from polarimetric radar data, replacing part of the
processing chain used in the Jacques et al. (2018) study.
While hoping that this replacement would bring an
improvement, this study also applies methods normally
used for the NWP model verification together with inde-
pendent observations as a means of establishing a frame-
work whereby improvements to data quality can be
objectively assessed and monitored over time. This ability
would effectively contribute towards an evidence-based
change management process, whereby decisions about
whether to introduce improvements to operational data
processing would be based on results from such scientific
methods.

2 | METHODS

The weather radar data used the study are US Next Gen-
eration Weather Radar (NEXRAD) Level II data from the
continental United States and Alaska, together with data
from the Canadian national network of 30 C-band Doppl-
er radars and the McGill University S-band polarimetric
radar outside Montréal, Québec, Canada. Together, the
total number of radars is about 185, depending on avail-
ability at any given moment. The study’s time period is
the complete month of August 2016.

2.1 | Radar processing

Jacques et al. (2018) document and reference a radar
data-processing chain for the purpose of obtaining a
radar product of sufficient quality for its assimilation. In
short, the chain comprises the following steps:
1. An ECCC in-house algorithm that combines non-Doppler and Doppler reflectivities, and performs vertical gradient testing to identify and remove clutter.

2. Hit-accumulation clutter filtering.

3. Identification of additional non-meteorological echoes, like as biological targets, radio interference and individual targets forming speckle.

4. Identification and correction of beam blockage by topography.

5. Conventional attenuation correction.

6. Descriptive algorithm describing radar beam broadening with increasing range.

The legacy ECCC C-band network is subject to the complete processing chain, the NEXRAD data are processed according to steps 3-6, and data from the McGill radar steps 4-6. A “total quality” indicator is also generated, being the minimum quality value resulting from the steps 4-6, and used for compositing. Jacques et al. (2018) can be consulted for more detail on both input data and processing.

The third step in this processing chain contains a combination of image analysis algorithms (Peura, 2002) for identifying and removing non-precipitation echoes from biometeor or aerocological targets (birds, insects), speckle, and external emitters (e.g. radio interference) in radar reflectivity data. These algorithms are replaced the present study with a method recently reported on by Kilambi et al. (2018) based on polarimetric information to perform roughly the equivalent task. This method was chosen herein because those authors presented and described it as being simple and robust, while it could perform on a par with more complex target classification algorithms. As the downstream use of the results only required knowledge about whether or not echoes were from precipitation, this simple approach implied a faster implementation and deployment in software, which was also deemed an advantage.

The new approach makes use of the DR (Melnikov and Matrosov, 2013; Ryzhkov et al., 2014, 2017), which is derived from a combination of differential reflectivity \(Z_{\text{DR}}\) and a co-polar correlation co-efficient \(\rho_{\text{HV}}\), and is referred to here as DRQC. The DR itself (Equation 1) describes how targets in each range bin are simultaneously uniform in shape and close to spherical (Kilambi et al., 2018). The authors of that paper illustrate how precipitation has a clear concentration of DR values coinciding with a combination of \(Z_{\text{DR}}\) near zero and \(\rho_{\text{HV}}\) near 1, whereas non-precipitation has no such DR signature.

They identify a DR threshold of \(-12\ dB\), below which targets are generally considered to be precipitation; reflectivity > 35 dBZ are exempted from this DR threshold and are preserved because such strong echoes are unlikely to be caused by biological targets. In fact, high DR associated with low \(\rho_{\text{HV}}\) and high reflectivity will be observed in the presence of graupel and hail, which should not be filtered out.

\[
\text{DR} = \frac{Z_{\text{DR}} + 1 - 2Z_{\text{DR}}^{1/2} \rho_{\text{HV}}}{Z_{\text{DR}} + 1 + 2Z_{\text{DR}}^{1/2} \rho_{\text{HV}}} \quad (1)
\]

Thresholding DR as described above is the first of two parts of the DRQC method, where the second part is a speckle filter. In polar space, each sweep of radar data is filtered using a five ray by seven bin kernel for each range bin of the Level II “super resolution” (0.5” by 250 m) data. In this kernel, the value in the majority or plurality among the following three categories is the winner: i) precipitation, ii) non-precipitation and iii) no echo. The first two categories are based on the DR and reflectivity thresholding, and the third category is a special value reserved to identify areas within radar coverage that have been radiated but which have not produced any echo. Values considered non-precipitation based on the output of this speckle filter are assigned this “no echo” special value.

No other changes were made to the radar data-processing chain presented by Jacques et al. (2018). The time period used in the present study took place before any of the new Canadian S-band radars, having the ability to make polarimetric measurements and allowing the application of the DRQC, had been deployed, so only the American data were processed with this algorithm out to a maximum range of 250 km. Kilambi et al. (2018) developed this approach with the NEXRAD Level II (S-band) data, and also applied it to the C- and X-bands data to demonstrate its transferability among wavelengths.

## 2.2 | Assessing data quality

Meteorological aerodrome reports (METARs) provide independent in situ surface-based observations that we have used to evaluate radar data quality and the impacts of the quality control methods. The METARs from 2,706 sites (Figure 1) in Canada and the continental United States are provided every hour valid on the hour. These observations contain valuable information on the occurrence of several different types of precipitation, and such observations are instantaneous in nature. Such precipitation occurrence observations can represent very weak precipitation intensities, which allows a relatively direct comparison with weather radars, which are very sensitive and can also detect very weak precipitation. Non-occurrence of precipitation is inferred from the METAR
present weather reports. Such reports have no explicit representation when there is no present weather observed and therefore actually no precipitation. In such cases, to make sure zero precipitation occurrence is only assigned to stations that normally report present weather, we check whether both horizontal visibility and cloud coverage are reported, because these two parameters are required to report explicitly in all the METARs no matter what is the weather (ICAO, 2018). Only then is a zero value assigned.

When co-locating each METAR with radar data, the nearest geometrical radar range bin in the lowest sweep of radar data in each input volume was used from all available radars at the top of every hour. This nearest radar bin is used, with no consideration to its neighbours, even if there is no echo in that bin. Any valid radar echo represents the presence of precipitation. Pairing the radar and METAR observations in this way brings the advantage that a given METAR can be located within several radars’ coverage areas, thereby giving a larger statistical sample of around 7,000 hourly, and 168,000 daily, comparisons in total.

An NWP model verification framework developed at ECCC, called Emet after the Hebrew word תמא for “truth”, allows for automated generation of standard statistical contingency table skill scores (Lemay and Husson, 2017). We have introduced precipitation occurrence as a standard verification variable in Emet. The skill scores of interest are summarized in Table 1. Such skill scores were aggregated into daily precipitation occurrence data from both the radar and METARs. Monthly skill scores were derived based on the daily results. To test whether the differences between skill scores using old and new quality controls are statistically significant, a three-day block bootstrapping method (Jolliffe, 2007) was used. This resampling method divides up the entire length of time (one month) into regular blocks, in this case three days each, giving 10 blocks. Doing this is an attempt to minimize correlation among the blocks containing skill score results. Such resampling is performed using data replacement, as opposed to data omission in cross-validation, to estimate the statistical properties for significance testing. In the present case, Emet resamples blocks 1,000 times to give reliable statistics with a 95% confidence interval.

Emet contains functionality for stratifying skill score statistics based on radar echo strength and other user-defined criteria based on the characteristics of the data. The present paper focuses on all (unthresholded) reflectivity data acquired by the radar, in the lowest polar sweep of data, up to a height of 5 km above the radar. No thresholding of the radar data implies no impacts on the radar’s sensitivity, thereby increasing the likelihood that real precipitation will be observed by both the radar and METAR. With a lowest elevation angle of 0.5°, this translates to a slant range of just over 225 km. The METARs from the contiguous United States, the lower 48 states,
are also only used because the differences in quality control using the DRQC only apply to radar data from the NEXRAD network.

2.3 | Computational procedures

The software environments used to generate the results in the present paper were collected into a project repository, publicly accessible at Michelson (2019). Part of this project is to preserve specific versions of the procedures used in an effort to achieve reproducibility according to Irving (2016).

The weather radar data-processing environment is the BALTRAD\textsuperscript{1} Toolbox (Michelson et al., 2018). This is one of several open-source software environments that have emerged in recent years supporting the weather radar community (Heistermann et al., 2014), even integrated together into working, publicly available solutions (Heistermann et al., 2015).

Figures were created using Matplotlib (Hunter, 2007), which is the standard plotting environment in the Python programming language. Scientific colour maps used with plotting radar data were based on those provided by Crameri (2018), with the goal of displaying relative differences uniformly with their continuous scales. Additionally, Pandas (McKinney, 2017) was used to prepare the plots containing skill score results.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure2.png}
\caption{One kilometre horizontal resolution Cartesian images from the new Canadian S-band radar at Radisson, Saskatchewan, on May 15, 2019, at 0900 UTC, 0300 LST: (a) differential reflectivity ($Z_{DR}$); (b) co-polar correlation co-efficient ($\rho_{HV}$); (c) radial wind velocity with two arrows indicating the direction of targets; and (d) derived depolarization ratio (DR)\label{fig:figure2}}
\end{figure}
3 | EXAMPLE OF RADAR DATA QUALITY CONTROL

Although Canadian S-band data from newly deployed dual-polarization radars were not available for the present study in 2016, such an example using recent data from Radisson, Saskatchewan, is presented, illustrating both old and new quality control procedures used in the third of the six steps outlined above in the radar data quality control processing chain. The case is from May 15, 2019, at 0900 UTC (0300 LST). What makes this case noteworthy is that it contains a large springtime bird migration, which both old and new quality control approaches are designed to identify and suppress, together with widespread storms that the methods are designed to preserve. There is no reliable reference map for this instantaneous case, as the radar is the only observing system in the area with high-resolution monitoring capabilities in both time and space.

Original polar data are \(0.5^\circ \times 500\) m resolution in this case, and the data were acquired out to a range of 240 km. Figure 2a–c shows 1 km pseudo-constant altitude plan position indicator pseudo-\((\text{CAPPI})\) images containing \(Z_{\text{DR}}\) and \(\rho_{\text{HV}}\), along with a lowest \((0.4^\circ)\) plan position indicator (PPI) containing radial wind velocity. Figure 2d contains a one km pseudo-CAPPI with derived DR, with the \(-12\) dB threshold generally delineating precipitation in shades of blue \((< -12\) dB) and non-precipitation in earth tones \((\geq -12\) dB). \(Z_{\text{DR}}\) is positive in general, with negative values apparent as a light blue speckle. A large region to the east of the radar shows relatively high values of \(Z_{\text{DR}}\) combined with relatively low \(\rho_{\text{HV}}\). The highest values of \(\rho_{\text{HV}}\) coincide in general with a near-zero \(Z_{\text{DR}}\), which is consistent with rain. Adding

**Figure 3** Horizontally polarized radar reflectivity from the same case as for Figure 2, with: (a) no quality control applied; (b) the old image analysis-based techniques applied; and (c) the new method applied based on depolarization ratio (DR)
additional value is information on radial wind velocities. If the general separation between precipitation and non-precipitation conveyed by the DR field is assumed to be valid, the winds appear to be southwesterly. There are exceptions, however, one being the area to the east of the radar that indicates southeasterly winds, which seems consistent with a biological target of migrating birds. This is also corroborated by the presence of two isodops, lines of equal zero Doppler velocity, each indicated by an arrow in Figure 2c: one in precipitation west of the radar, the other in non-precipitation (birds) to the north of the radar shifted by roughly 90°.

The horizontally-polarized reflectivity field is illustrated in Figure 3. The old quality control methods (Figure 3b) succeed in identifying and removing a significant amount of biological targets to the north and east of the radar, between areas of precipitation. They also succeed in removing many of the biological echoes in the southern half of the radar’s coverage area. However, they have difficulty in preserving what can be assumed to be real precipitation in some of these same areas, while also failing to identify and remove non-precipitation targets in areas to the north and east that are suspected to be biological. Beyond around 190 km in range, only small speckle is removed, which has the effect of preserving that we suspect is precipitation. In general, what is preserved has the appearance of an area of widespread storms associated with the passing of a warm front, if only the reflectivity field is considered.

Figure 3c shows the results of the quality control based on DR. Unsurprisingly, the areas illustrated in Figure 2d are assumed to be of precipitation remain largely preserved, whereas the non-precipitation has been largely suppressed. In this case, the speckle filter has used a polar kernel of five rays by five range bins, owing to the lower resolution in range compared with the NEXRAD data. The speckle filter has suppressed small-scale precipitation, representing areas of nine range bins large or smaller, at all ranges, but it also succeeded in keeping intact larger precipitation areas.

This example shows the strengths and weaknesses of both quality control approaches in non-trivial conditions, highlighting the difficulty in removing unwanted information while preserving what is wanted. What looks encouraging in the application of the new method needs to be objectively evaluated to determine whether improvements to data quality when using it are real or just perceived.

It should also be noted that while this case helps illustrate the behaviour of our data and both old and new quality control methods, the new Canadian radar configuration and scan strategy are still undergoing assessment with the objective of identifying improvements, implying that changes may be forthcoming.

4 | RESULTS

Average skill scores aggregated from hourly scores from each day of August 2016 are illustrated in Figures 4–6. Summary statistics for the whole month are given in Table 2. These results show several noteworthy features. The probability of detection (POD) from un-quality controlled radar data cannot be improved upon using any of the methods applied in the present study, but there is a large improvement in the POD when using the new radar quality control methods compared with the old ones. The use of both old and new radar quality control methods lower the false-alarm ratio (FAR) greatly, with the new quality controls performing slightly better visually than the old ones (Figure 4).

The Peirce skill score (PSS) and Heidke skill score (HSS) both identify characteristics of overall skill that complement each other, in that the PSS describes the separability of the radar and METAR precipitation occurrences, whereas the HSS describes the separability between radar-based precipitation occurrence and the chance that precipitation would occur randomly. According to the PSS, un-quality controlled radar data
have higher skill than radar data quality controlled using the old methods. This condition is reversed with the new radar quality controls. A benchmark HSS value in the interval 0.3–0.4 is considered to represent a minimum meaningful skill. Radar data that have not been quality controlled are largely below this interval. The old radar quality controls perform better, averaging just under 0.4 for the month. The new radar quality controls give HSS values representing consistently meaningful skill, with an average > 0.5 (Figure 5).

The frequency bias index (FBI) indicates to what extent the radar data are over- or under-observing the occurrence of precipitation relative to the reference given by the METAR-based precipitation occurrence observations. Radar data that are not quality controlled clearly over-observe the occurrence of precipitation, and this should come as no surprise since the objective is to suppress the non-meteorological content in the radar data. The old quality controls yield daily FBI values systematically below 1 and as low as just below 0.3, indicating that a significant amount of precipitation is being removed in addition to the non-precipitation. The new radar quality controls give FBI values much closer to 1, indicating higher skill with this metric (Figure 6).

With the exception of POD, the radar data processed using the new quality controls have consistently improved skill compared with radar data that have either not been quality controlled or quality controlled using the old methods. All differences between skill scores using old and new quality controls were statistically significant using a 95% confidence interval.

5 | DISCUSSION AND CONCLUSIONS

The application of the dual-polarization-based quality control method that combines the DR and a speckle filter (Kilambi et al., 2018) yields higher quality continental American radar data compared with legacy image-analysis methods based on reflectivity only. This has been objectively determined through independent precipitation occurrence observations given by hourly METARs.

There is nothing new about objectively determining forecast skill or data quality using independent observations and skill scores such as those presented herein. However, the way we have organized our workflow has been with the ambition to create a continuous monitoring of weather radar data quality, both before and following quality control, such that the impacts to data quality can be assessed over time. Doing so enables, in principle, the ability to reveal improvements to data quality,
brought about through changes to radar configuration and scan strategy, before quality control. Doing so also enables the introduction of improved quality control methods that can be better than the current ones, and determining the impacts to data quality when applying them. Monitoring data quality has been reported on previously, for example, by Frech (2013) who uses a suite of methods to monitor system performance and calibration stability of the German weather radar network continuously. We have extended this to downstream quality control, giving us a quality assurance process for radar data processing.

Precipitation occurrence is a blunt binary parameter, and the use of the METARs from across the North American continent is from a relatively low-density observation network. With 185 weather radars, the number of METARs per radar averages to under 40, but with great variability. While this is a reasonable number of comparisons for a single radar’s coverage area, not a lot of detail can be revealed. The strength in assessing the impacts of quality control at the continental scale is in the total amount of comparative data, and the authors are confident in the skill score results from one month of comparisons.

As the name suggests, the METAR sites are commonly at airfields. This would seem to indicate areas with relatively flat terrain and therefore not prone to contamination by clutter in radar data, thereby making them unrepresentative in general. However, airports contain concentrations of infrastructure that are good clutter targets: objects with large radar cross-sections such as buildings, towers and aircraft, so it is reasonable to expect that many such sites would be suitable for testing the effectiveness of quality control. Another feature in our skill scores worth clarifying is the relatively low values of POD, averaging around 0.6 for data that have not been quality controlled (Figure 4a and Table 2). As mentioned in Section 2.2, radar data from the lowest sweep (0.5°) were used up to a height of 5km above the radar, corresponding to a range of around 225 km. This is relatively distant from the radar, and the low POD values represent precipitation occurrences at the surface that are undetected by the radar aloft, which is an effect of normal weather radar observing geometry. Reducing the height of detection used to calculate the skill scores would have the effect of reducing the radars’ effective coverage areas, thereby drastically reducing the statistical sample sizes, which we wanted to avoid.

Further exploitation of the information content in the METARs, such as precipitation type, might be worth exploring, as would further statistical stratification, for example regionally, to reveal the strengths and weaknesses of quality control. Other independent observations and skill scores would offer insights that would complement the information gained from the METARs. Stratification of skill based on imposing a reflectivity threshold (reduced radar sensitivity) is also a possibility, although doing so would be unwise because it could potentially reduce the number of precipitation events observed by both the radar and METAR. A follow-up study could,

| Skill score | QC  | Minimum | Maximum | Mean  | Standard deviation |
|-------------|-----|---------|---------|-------|-------------------|
| POD         | None| 0.474   | 0.725   | 0.605 | 0.055             |
| POD         | Old | 0.229   | 0.428   | 0.319 | 0.052             |
| POD         | New | 0.443   | 0.679   | 0.553 | 0.054             |
| FAR         | None| 0.640   | 0.928   | 0.815 | 0.070             |
| FAR         | Old | 0.289   | 0.591   | 0.439 | 0.077             |
| FAR         | New | 0.303   | 0.528   | 0.427 | 0.059             |
| PSS         | None| 0.362   | 0.577   | 0.462 | 0.052             |
| PSS         | Old | 0.220   | 0.409   | 0.306 | 0.051             |
| PSS         | New | 0.425   | 0.648   | 0.532 | 0.051             |
| HSS         | None| 0.100   | 0.388   | 0.218 | 0.073             |
| HSS         | Old | 0.301   | 0.492   | 0.381 | 0.050             |
| HSS         | New | 0.461   | 0.655   | 0.538 | 0.045             |
| FBI         | None| 1.885   | 8.019   | 3.765 | 1.423             |
| FBI         | Old | 0.368   | 0.771   | 0.576 | 0.103             |
| FBI         | New | 0.766   | 1.135   | 0.972 | 0.094             |
however, include range from the radar as a variable in order to determine the sensitivity of the radar at detecting precipitation occurrence observed by the METARs.

The DR is based in part on $Z_{DR}$, which is well known to be prone to bias. While it is desirable to know and correct for such biases, neither Kilambi et al. (2018) nor we have done so. In their study they both note and illustrate why DR is less sensitive to $Z_{DR}$ bias than fuzzy logic-based target classification approaches. This robustness has not yet been quantified, and doing so would be a useful exercise in a follow-up study, assuming the availability of reliable $Z_{DR}$ offsets, which we have not had in this study.

A natural next step is to assess the improved quality of the data further through a radar data assimilation impact study such as that reported by Jacques et al. (2018). Other downstream applications of these data sets, such as quantitative precipitation estimation and radar-based precipitation nowcasting, stand to benefit from improved reliability, and that the echoes represent precipitation, in the quality-controlled radar data resulting from the use of DRQC. Before applying the data in such applications, other quality controls continue to be required to address issues such as beam blockage from topography and attenuation from precipitation. Knowing that such methods are not applied to non-precipitation echoes implies that their use is becoming more physically meaningful, and reduces uncertainty in downstream application of the data.

The data used for the study derived from summer conditions, where the DRQC method is expected to work best. With cooler conditions in autumn come well-developed melting layers (bright bands) at closer ranges to the radars, which bring significantly lower values of $\rho_{HV}$ and higher DR which will likely lead to the removal of some real precipitation in these areas. In turn, this will likely worsen the skill scores, indicating the lower effectiveness of DRQC and lower radar data quality. The impacts of the DRQC approach will require monitoring in different seasons, and the knowledge collected is expected to guide the authors in either evolving the approach or choosing another. A follow-up to the present study could look at how the DRQC performs through a complete transition between summer and winter conditions, a period of around five to six months, which we have been unable to do in this initial study. It can be speculated that well-developed winter conditions, that is, near-freezing or sub-zero surface temperatures with or without precipitation, would have a much lower proportion of non-precipitation echoes, leading to the lowered effectiveness of the method but with high-quality data resulting anyway.

The DRQC approach presented in the paper has been running operationally at Environment and Climate Change Canada (ECCC) since October 2019. Additional improvements to data quality will be determined through the same assessment process as presented herein, and decisions made on deploying them operationally will be based on the results. The change management process is thereby based on scientific evidence.

ACKNOWLEDGEMENT

Professor Frédéric Fabry of McGill University is gratefully acknowledged for introducing us authors to the approach of quality control using depolarization ratio.

ORCID

Daniel Michelson https://orcid.org/0000-0003-3222-9881

ENDNOTE

1Not an acronym, BALTRAD was developed to network weather radar data in Europe’s Baltic Sea region.

REFERENCES

Alberoni, P.P., Andersson, T., Mezzasalma, P., Michelson, D.B. and Nanni, S. (2001) Use of the vertical reflectivity profile for identification of anomalous propagation. Meteorological Applications, 8, 257–266. https://doi.org/10.1017/S135048270003012.

Bech, J., Gjertsen, U. and Haase, G. (2007) Modelling weather radar beam propagation and topographical blockage at northern high latitudes. Quarterly Journal of the Royal Meteorological Society, 133, 1191–1204. https://doi.org/10.1002/qj.98.

Cramer, F. (2018) Scientific colour-maps. Zenodo. http://doi.org/10.5281/zenodo.1243862.

Donaldson, R.J., Dyer, R.M. and Kraus, M.J.. (1975) An objective evaluator of techniques for predicting severe weather events. Preprints, Ninth Conference on Severe Local Storms. Norman, Oklahoma: American Meteorological Society, pp. 321–326

Fabry, F. (2015) Radar Meteorology: Principles and Practice. Cambridge University Press ISBN 9781108460392.

Frech, M. (2013) Monitoring the data quality of the new polarimetric weather radar network of the German Meteorological Service. Proceedings of the AMS 36th Conference on Radar Meteorology. Available at: https://ams.confex.com/ams/36Radar/webprogram/Manuscript/Paper228472/ams_radar2013_frech_monitoring_9_B3.pdf.

Fulton, R.A., Breidenbach, J.P., Seo, D.-J. and Miller, D. (1998) The WSR-88D rainfall algorithm. Weather and Forecasting, 13, 337–395. https://doi.org/10.1175/1520-0434(1998)013<0377:TWRA>2.0.CO;2.

Germann, U. and Joss, J. (2004) Operational measurement of precipitation in mountainous terrain. Chapter 2. In: Meischnner, P. (Ed.) Weather Radar. Berlin Heidelberg: Springer Verlag, pp. 52–77. https://doi.org/10.1007/978-3-662-05202-0_2.

Gourley, J.J., Tabary, P. and Parent du Chatelet, J. (2007) A fuzzy logic algorithm for the separation of precipitating from non-precipitating echoes using polarimetric radar observations. Journal of Atmospheric and Oceanic Technology, 24, 1439–1451. https://doi.org/10.1175/JTECH2035.1.
Ryzhkov, A., Matrosov, S.Y., Melnikov, V., Zrnic, D., Zhang, P., Cao, Q., Knight, M., Simmer, C. and Troemel, S. (2017) Estimation of depolarization ratio using weather radars with simultaneous transmission/reception. *Journal of Applied Meteorology and Climatology*, 56, 1797–1816.

Saltikoff, E., Cho, J.Y.N., Tristant, P., Huuskonen, A., Allmon, L., Cook, R., Becker, E. and Joe, P. (2016) The threat to weather radars by wireless technology. *Bulletin of the American Meteorological Society*, 97, 1159–1167. https://doi.org/10.1175/BAMS-D-15-00048.1.

Vivekanandan, J., Zrnic, D.S., Ellis, S.M., Oye, R., Ryzhkov, A.V. and Straka, J. (1999) Cloud microphysics retrieval using S-band dual-polarization radar measurements. *Bulletin of the American Meteorological Society*, 80(3), 381–388. https://doi.org/10.1175/1520-0477(1999)080<0381:CMRUSB>2.0.CO;2.

Wessels, H. and Beekhuis, H. (1994) Stepwise procedure for suppression of anomalous ground clutter. In: *COST-75 Seminar on Advanced Radar Systems*. Brussels: European Commission, EUR 16013 EN, pp. 270–277.

Ye, B.-Y., Lee, G.W. and Park, H.-M. (2015) Identification and removal of non-meteorological echoes in dual-polarization radar data based on a fuzzy logic algorithm. *Advances in Atmospheric Sciences*, 32(9), 1217–1230. https://doi.org/10.1007/s00376-015-4092-0.

Zrnić, D.S. and Ryzhkov, A.V. (1999) Polarimetry for weather surveillance radars. *Bulletin of the American Meteorological Society*, 80, 389–406. https://doi.org/10.1175/1520-0477(1999)080<0389:PFWSR>2.0.CO;2.

How to cite this article: Michelson D, Hansen B, Jacques D, Lemay F, Rodriguez P. Monitoring the impacts of weather radar data quality control for quantitative application at the continental scale. *Meteorol Appl.* 2020;27:e1929. https://doi.org/10.1002/met.1929