

A Centimeter-Wavelength Snowfall Retrieval Algorithm Using Machine Learning

FRASER KING,^a GEORGE DUFFY,^{b,c} AND CHRISTOPHER G. FLETCHER^a

^a *Department of Geography and Environmental Management, University of Waterloo, Ontario, Canada*

^b *NASA Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California*

^c *Earth and Environmental Sciences, Syracuse University, Syracuse, New York*

(Manuscript received 14 March 2022, in final form 26 May 2022)

ABSTRACT: Remote sensing snowfall retrievals are powerful tools for advancing our understanding of global snow accumulation patterns. However, current satellite-based snowfall retrievals rely on assumptions about snowfall particle shape, size, and distribution that contribute to uncertainty and biases in their estimates. Vertical radar reflectivity profiles provided by the vertically pointing X-band radar (VertiX) instrument in Egbert, Ontario, Canada, are compared with in situ surface snow accumulation measurements from January to March 2012 as a part of the Global Precipitation Measurement (GPM) Cold Season Precipitation Experiment (GCPEX). In this work, we train a random forest (RF) machine learning model on VertiX radar profiles and ERA5 atmospheric temperature estimates to derive a surface snow accumulation regression model. Using event-based training–testing sets, the RF model demonstrates high predictive skill in estimating surface snow accumulation at 5-min intervals with a low mean-square error of approximately $1.8 \times 10^{-3} \text{ mm}^2$ when compared with collocated in situ measurements. The machine learning model outperformed other common radar-based snowfall retrievals (Z_e – S relationships) that were unable to accurately capture the magnitudes of peaks and troughs in observed snow accumulation. The RF model also displayed consistent skill when applied to unseen data at a separate experimental site in South Korea. An estimate of predictor importance from the RF model reveals that combinations of multiple reflectivity measurement bins in the boundary layer below 2 km were the most significant features in predicting snow accumulation. Nonlinear machine learning–based retrievals like those explored in this work can offer new, important insights into global snow accumulation patterns and overcome traditional challenges resulting from sparse in situ observational networks.


KEYWORDS: Snowfall; Remote sensing; Machine learning

1. Introduction

There are many reasons why it is important to measure falling snow. Accumulated snow has important linkages to regional flooding, water resource management practices, and ecosystem development and sustainability (Buttle et al. 2016; Berghuijs et al. 2019; Déry and Brown 2007; Peng et al. 2010). As global average temperatures continue to rise, the distribution and magnitude of synoptic snowfall events are expected to also change in a substantial, nonlinear manner (Vaughan et al. 2013). In many cases, the most effective way to measure snow is through radar. Satellite radar measurements, for example, can provide valuable snowfall information in remote regions such as the high Arctic or tundra, or over sea ice (King 2019; Cabaj et al. 2020; Duffy et al. 2021). They can also be used to retrieve measurements of melting snow in the ocean that may represent up to 10% of the seasonal ocean surface heat flux (Duffy and Bennartz 2018). Satellite-based snowfall radar retrievals have also been used as independent observational datasets to evaluate and constrain current gridded snow products across Arctic regions, which are typically poorly observed due to sparsely distributed in situ measurement networks (King and Fletcher 2021).

Radars are valuable resources as they can provide precipitation rates over a wider area and at a finer temporal scale relative to traditional in situ precipitation gauges (Lemonnier et al. 2019). Radars can only provide estimates of precipitation, however, and the accuracy of retrieved precipitation rates is largely dependent on the accuracy of the retrieval algorithm (Kulie and Bennartz 2009; Hiley et al. 2010). There are several common methods typically used to convert radar measurements into precipitation rates. So-called Z_e – S relationships (i.e., $Z_e = aS^b$) use algebraic power laws, with the prefactor a and exponent b derived from experimental observations or simulated models to convert near-surface reflectivity Z_e into precipitation rates S (Wolfe and Snider 2012; Boucher and Wieler 1985; Szyrmer and Zawadzki 2010). Bayesian methods, such as the techniques used in the derivation of the 2C-SNOW-PROFILE snowfall retrieval product from *CloudSat* (Stephens 2017), are also used for snowfall retrieval by incorporating prior assumptions of snowfall particle size, shape, distribution, and fall speed (Wood and L'Ecuyer 2021).

These empirically derived relationships may only provide accurate estimates under specific climate regimes as a consequence of the necessary a priori assumptions made about particle microphysics. Substantial error can therefore propagate through to the final snow accumulation estimates if the wrong Z_e – S relationship is applied (Milani et al. 2018; von Lerber et al. 2017; Schoger et al. 2021). The variability in the structure and vertical extent of winter storms can also lead to additional uncertainty and error when deriving a surface snow accumulation estimate from only near-surface reflectivity bins when applying such a relationship. While the significance of each

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

Corresponding author: Fraser King, fdmking@uwaterloo.ca

radar bin may vary, an ideal retrieval algorithm would incorporate all of the relevant reflectivity information that a radar beam can provide to maximize the accuracy of the estimated surface accumulation. Such an algorithm would also be able to assimilate cloud features and environmental conditions that are known to further influence Z_e - S relationships and Bayesian retrievals (Pettersen et al. 2020; McIlhatten et al. 2020).

Traditional algebraic methods for precipitation retrievals can become impractical if we consider the full set of variables present in a typical radar retrieval. Machine learning (ML) algorithms, however, are unrestricted by algebra and can incorporate a virtually unlimited number of variables to derive complex and nonintuitive relationships between the provided predictors and response (Karpatne et al. 2019). Previous studies have used ML to construct snowfall retrieval algorithms that can detect and estimate surface snowfall rates under specified climate contexts using spaceborne passive microwave radar measurements (Adhikari et al. 2020; Ehsani et al. 2021). Neural networks have also recently shown promise in retrieving particle size distribution parameters from synthetic model data that, when applied to spaceborne radar observations, demonstrates improved estimates of snowfall relative to their default algorithms (Chase et al. 2021). These models are powerful tools for deriving snowfall estimates from radar data, and in further advancing our understanding of how accumulated snow quantities are changing under a warming global climate.

Herein, we develop and evaluate a novel random forest (RF) snow accumulation retrieval algorithm that incorporates reflectivity and temperature from all applicable radar range gates. The RF model is then compared with traditional Z_e - S relationships (derived from X-band radar) to evaluate model skill. Then, both the RF and Z_e - S relationships are applied to data from a completely unseen site in South Korea to evaluate model robustness. We use the RF algorithm to identify which radar gates appear as the most important, along with how many gates are necessary to maximize the accuracy and robustness of the retrieval. For the purposes of this methodological proof-of-concept study in the development of a retrieval strategy that could be applied to terrestrial and spaceborne radar, we are focusing on X-band radar measurements of snowfall to minimize potential complications that may arise from attenuation, non-Rayleigh reflectivity, or multiple scattering.

The main objectives of this paper are to

- 1) quantify the accuracy with which we can model surface snow accumulation from X-band radar data using an ensemble of statistical techniques of varying complexity,
- 2) determine whether ML models offer improved skill relative to traditional Z_e - S relationships, and
- 3) identify the most significant range gate estimates of temperature and reflectivity to enhance current and future retrievals of surface snow accumulation.

2. Data and methods

a. *In situ* snow observations

In situ precipitation measurements were obtained from the Global Precipitation Measurement (GPM) Cold Season

Precipitation Experiment (GCPEX) field campaign (Skofronick-Jackson et al. 2015). GCPEX was a joint venture between the National Aeronautics and Space Administration (NASA) and Environment and Climate Change Canada. A primary goal of GCPEX was to provide multiple observations of snowstorms in order to evaluate GPM-style retrievals of precipitating clouds. GCPEX took place across several sites in Ontario, though for this study we focus on observations taken at the Centre for Atmospheric Research Experiments in Egbert, Ontario, Canada, from January to March 2012 (Fig. 1a).

Surface snow accumulation was monitored using an OTT Pluvio2 automated precipitation weighing gauge at the GCPEX study site. The gauge was installed at a representative location in an open field, free from the influence of human activity and the presence of nearby vegetation. Snow accumulation estimates were recorded at 1-min temporal resolution to provide $n = 25$ days of nonzero precipitation observations. The Pluvio2 minimum observational intensity threshold is 0.2 mm min^{-1} of snow accumulation for measurements to qualify as active precipitation and be registered by the system (i.e., trace amounts of accumulation). Accumulation estimates that fell below this minimum threshold were not considered as precipitation and therefore left unrecorded. This measurement gauge was equipped with a 200-cm^2 heated top to help to prevent snow accumulation from covering the gauge top orifice and mitigate against underestimation. The Pluvio2 was also equipped with 8-ft-diameter (244 cm) Alter shield and was situated in the center of a surrounding Tretyakov fence (Smith 2009; Metcalfe et al. 1997). This double-fence setup helps to mitigate the effects of turbulent winds that can deflect incoming hydrometeors away from the gauge-top opening and lead to an underestimation of accumulation (i.e., undercatch effects) (Sevruk et al. 1991; Kochendorfer et al. 2017; Rasmussen and Rasmussen 2014; Colli et al. 2020). To further mitigate against issues with strong winds and gauge undercatch, we constrain our analysis to periods where surface wind speeds recorded during GCPEX were less than 5 m s^{-1} (i.e., only low-to-medium intensity wind speed periods) for the duration of the measurement period (Kochendorfer et al. 2022; Rasmussen et al. 2002).

Additional surface snow accumulation observations and collocated X-band radar data were collected from the International Collaborative Experiment for PyeongChang Olympic and Paralympics (ICE-POP) in South Korea (Fig. 1c). Similar to GCPEX, ICE-POP was another GPM Ground Validation campaign that collected a suite of atmospheric observations for researching frozen and mixed-phase precipitation (Gatlin 2020). The experiment included a Pluvio2 gauge with a similar double-fenced shield to that at GCPEX, as well as a collocated MXPOL X-band radar system (Gehring et al. 2021). Because the MXPOL data extend beyond 10 km, the lowest subset of the vertical column is extracted and aligned to the vertical resolution (44.5 m) of bins used at GCPEX (Fig. 1d). Preprocessing of the MXPOL and Pluvio2 data from ICE-POP follows the same steps as those applied to data at GCPEX (further details are in section 2d).

To address uncertainties arising from trace amounts of accumulation being missed at the surface Pluvio2 gauge, we examined varying levels of temporal data averaging and also

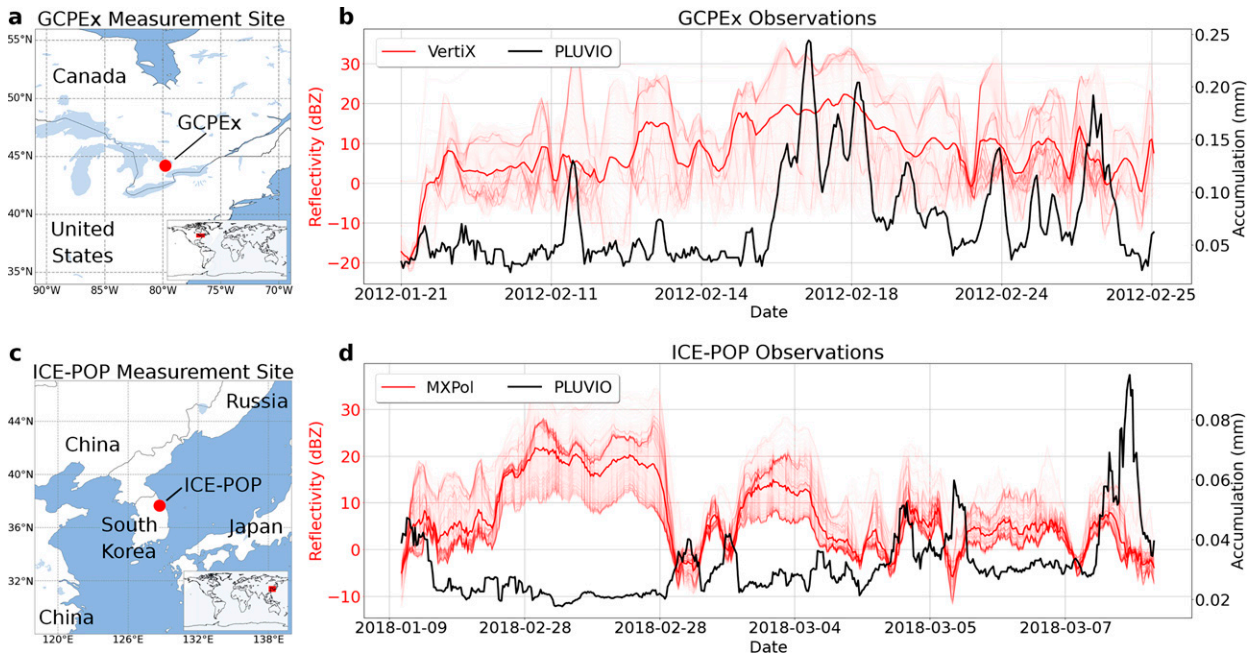


FIG. 1. (a) The measurement site location at Egbert, (b) 5-min VertiX reflectivity values (mean in thick red and individual vertical bins in light red) and Pluvio2 snow accumulation estimates (black) from the site in (a), (c) the ICE-POP measurement site location in South Korea, and (d) as in (b), but for ICE-POP using MXPOL X-band radar observations.

performed sensitivity analyses on minimum precipitation thresholds to determine whether certain low-intensity observations should be included in the model training data. We examined minimum accumulation thresholds from 0.01 to 0.16 mm and found that each model demonstrated similar strong performance (low MSE) for minimum thresholds of 0.01 and 0.02 mm, with reductions in performance as we move to larger minimum thresholds. For example, we note a 64% increase in MSE for RF_{full} trained using GCPEX data and applied at ICE-POP when the training threshold is increased from 0.01 to 0.04 mm (and a 175% increase in MSE when the training threshold is increased from 0.01 to 0.16 mm). To maintain high accuracy and a large sample size, we therefore employ a minimum 5-min-average accumulation threshold of 0.01 mm.

b. VertiX radar

The vertically pointing X-band ($\lambda \approx 3$ -cm wavelength) radar (VertiX) provides reflectivity profiles with a sampling frequency of 3 s, a range of 10 km with vertical bin resolution of approximately 44.5 m, and a sensitivity below -20 dBZ. The VertiX instrument emits pulses of energy that scatter off of the particles within the cloud to provide precipitation echo measurements that can then be used to derive information about precipitating hydrometeors contained within the vertical column above the instrument (Skofronick-Jackson et al. 2015). This type of X-band radar has been used to great effect in previous work by Gehring et al. (2021) during experiments focused on analyzing precipitation type as well as accumulation quantification. This instrument was positioned in the same study location as the

Pluvio2 weighing gauge described in section 2a to provide time-height estimates of reflectivity (dBZ) and Doppler velocity ($m s^{-1}$) for precipitating clouds directly overhead both instruments. Reflectivity values and Pluvio2 accumulation estimates are shown for the full study site in Fig. 1b. While X-band radars have traditionally been used for rainfall retrieval algorithms, recent studies suggest that this radar band can provide valuable information related to snowfall rates in cold environments, and the longer wavelength relative to Ku/Ka radar helps to mitigate against the effects of non-Rayleigh scattering of large snow particles (Matrosov et al. 2009, 2007).

c. ERA5 reanalysis

Atmospheric temperature data are provided by the fifth major global reanalysis produced by the European Centre for Medium-Range Weather Forecasts (ERA5) dataset. ERA5 is a global reanalysis product that provides hourly estimates of a suite of climatic variables at $0.25^\circ \times 0.25^\circ$ resolution from 1979 to the present, with vertical column estimates up to 80 km into the upper atmosphere (Hersbach et al. 2020). As a fifth-generation product, ERA5 is well documented, and its estimates of temperature have been extensively analyzed and validated in previous literature (Tarek et al. 2020; Betts et al. 2019; Simmons et al. 2020; Wang et al. 2019). Hourly ERA5 atmospheric temperature data were extracted for each bin in the lowest 10 km of its vertical profile for the grid cell directly overlapping with the Egbert study site over our study period. Vertical bins in ERA5 were then regridded to match the approximate 44.5 m vertical bins defined by the VertiX radar system. The resulting binned temperature data

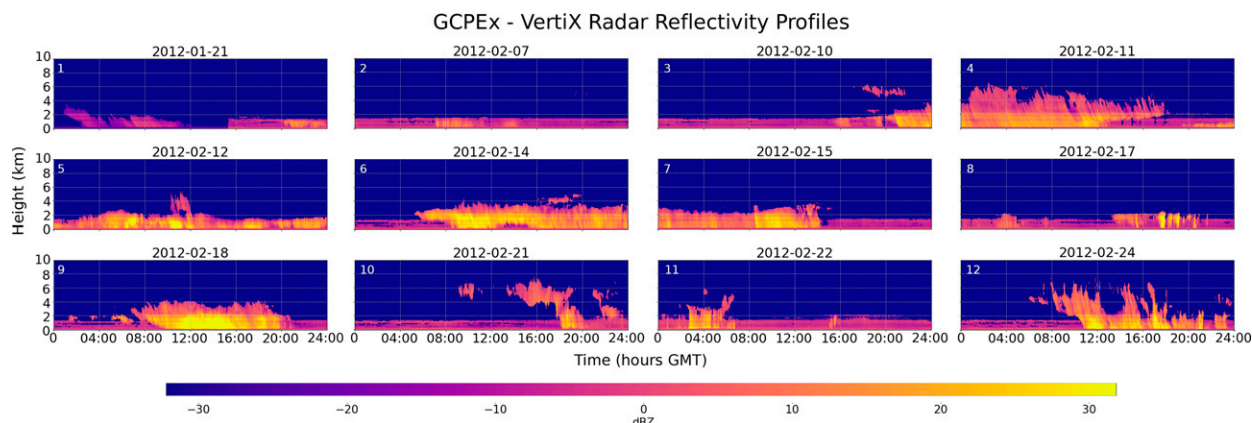


FIG. 2. Vertical radar reflectivity profiles generated from the VertiX radar over $n = 12$ collocated precipitation events (labeled by number in the upper left of each panel).

are then used as an additional predictor set in the snowfall retrieval techniques described in section 2d.

d. Snow accumulation retrieval methodology

There were 12 days during GCPEX where snowfall was measured by both the VertiX and Pluvio2 instruments. The VertiX reflectivity profiles for these 12 days of collocated measurements are shown in Fig. 2 along with the date of occurrence. We note the unique structure of precipitation in each day of Fig. 2, along with a relatively uniform low-intensity band of reflectivity values (approximately equal to 0 dBZ on average), in the 0–1.5-km portion of the vertical profile. The impact of this near-surface reflectivity band on model skill is discussed in more detail in section 4.

To begin the comparison process, each vertical bin of reflectivities in the VertiX radar profile (in physical units of millimeters to the sixth power per meter cubed), along with each 1-min in situ accumulation estimate is temporally averaged to common 5-min intervals. This interval size was selected to help account for small jumps in trace amounts of precipitation from the in situ gauge and form a more continuous reference dataset while still maintaining a sufficient sample size for model training and testing ($n = 391$ observations). Furthermore, as only periods of nonzero precipitation were reported by the in situ gauge, our comparisons are restricted to periods of active snowfall.

Model performance is evaluated using mean-squared error (MSE) calculated from a nonshuffled 10-fold cross validation (CV) so as to use the full set of available data and to mitigate against potential model overfitting. MSE is a common metric

used in ML precipitation retrieval studies to quantify the average error between observed and predicted values and assess model skill (Chen et al. 2020; Kim and Bae 2017; Shi et al. 2017). The nonshuffled nature of each 90/10 train/test split allows for the model's predictive skill to be examined when tested on unseen data, because retrievals taken close in time during training may be autocorrelated and lead to overfitting (i.e., a shuffled-split approach).

Each RF retrieval model examined here is provided with the same set of training data containing predictors and response variables. The response variable is the in situ snow accumulation observations provided by the Pluvio2 instrument. The default model predictors include the 221 vertical reflectivity bins from the VertiX radar and the same 221 atmospheric temperature bins from ERA5. Model predictors can also take the form of aggregations and subsets of these bins depending on the model and experiment being performed. More information on the individual structures of these experiments can be found in Table 1, along with their general performance in section 3.

To assess RF model accuracy, a set of five Z_e - S relationships derived from X-band radar experiments were selected (Table 2) to act as baseline comparisons. These relationships were applied to the 11th temporally averaged radar bin from the surface as this region of the profile demonstrated strong predictive skill in our later analysis of vertical profile feature importance (section 3b) and lay far enough above the surface to avoid concerns of ground clutter contamination. Each relationship's surface snow accumulation estimate was then recorded for comparison with the RF over a common 5-min time step.

A set of bespoke Z_e - S relationships were also empirically derived using the 5-min-average reflectivity observations ($\text{mm}^6 \text{m}^{-3}$)

TABLE 1. Model list and summary descriptions of the predictors used within each model.

Model	Variables	Inputs	Description
RF _{full}	Reflectivity; temperature	442	All reflectivity and temperature bins
RF _{near}	Reflectivity; Temperature	90	Near-surface reflectivity and temperature bins
RF _{far}	Reflectivity; Temperature	352	Upper reflectivity and temperature bins
Z_e - S	Reflectivity	1	Z_e - S relations described in more detail in Table 2

TABLE 2. K-band precipitation power-law relationships derived from previous literature.

Name	Power law	Source
Dumont d'Urville (DDU)	$Z_e = 76 \times S^{0.91}$	Grazioli et al. (2017)
Braham (1990) A (B90A)	$Z_e = 67 \times S^{1.28}$	Matrosov et al. (2009)
Woods et al. (2008) A (W08A)	$Z_e = 28 \times S^{1.44}$	Matrosov et al. (2009)
Woods et al. (2008) B (W08B)	$Z_e = 36 \times S^{1.56}$	Matrosov et al. (2009)
Woods et al. (2008) C (W08C)	$Z_e = 48 \times S^{1.45}$	Matrosov et al. (2009)

and temporally aligned in situ snow accumulation measurements from the Pluvio2 at GCPEX and ICE-POP. These fits were performed under the same 10-fold CV strategy as the RF, using the Python SciPy package's *curve_fit* optimization algorithm to derive custom GCPEX and ICE-POP power-law relations denoted as GZS and IPZS, respectively. Derivations were performed using a nonlinear least squares approach for finding optimal a and b coefficients to fit a Z_e - S power-law relationship (Virtanen et al. 2020). GZS and IPZS were then applied to unseen reflectivity observations in the testing set (in the same manner as the other Z_e - S relationships) to derive corresponding surface snow accumulation estimates. GZS and IPZS were included in this work to act as additional baseline comparisons and to further demonstrate the robustness of a more sophisticated ML-based algorithm over traditional power-law relationships when derived from the same sets of site-specific observations.

The RF regression model is implemented using Python's scikit-learn package from Pedregosa et al. (2011), which provides access to a wide library of model setup, prediction, and evaluation functions. RF model hyperparameterization is performed using a random search 10-fold cross validation, where subsets of model hyperparameters are randomly tested across the full parameter space to identify a theoretical optimal set of values for our model and data. Using this technique, we examined the parameter space for the hyperparameters listed in Table 3 to identify optimal values.

3. Results

a. Model intercomparisons

Cumulative accumulation time series are shown for all models (colored lines) and Pluvio2 snow measurements (black dashed line) in Fig. 3a for the full study period. While all models display similar accumulation trends over the full study period, the RF appears to be the most capable in capturing peaks and troughs in the data over time to more closely model the correct magnitude of the total recorded surface snow accumulation. W08A

also closely models the total accumulation over the full period, but it (along with the other power-law relationships) displays a strong negative bias during the first half of the study. W08A and DDU then display a positive total accumulation bias during the 18 February snowfall event; however, all Z_e - S relationships (excluding GZS) underestimate the precipitation rate following this event (as the slope of the accumulation curves appear to flatten out relative to in situ). While GZS closely follows in situ accumulation during the trace precipitation events at the beginning of GCPEX, on February 10 it begins to overestimate the amount of surface accumulation and remains positively biased by two orders of magnitude more than RF_{full} for the remainder of the period when compared with in situ. We note that the RF displays the lowest overall error in its predictions with MSE values smaller than any of the other Z - S relationships.

As mentioned in section 1, the referenced Z_e - S power-law coefficients are empirically derived from a set of observed atmospheric particle microphysics. These assumptions about particle microphysics are a large source of uncertainty in retrieval accuracy if the snowfall event being observed falls outside measured values recorded at the experiment from which the power law was derived. These uncertainties contribute to additional retrieval error when applied to data outside the training dataset, which is likely a large source of the error between each Z_e - S relation and the Pluvio2 estimates.

To further examine model performance on unseen data and assess the robustness of the RF retrieval accuracy, we performed a secondary test against a similar experimental setup at the ICE-POP site in South Korea (Fig. 4a). Testing RF_{full} on ICE-POP data (while trained only on data from GCPEX) allows us to examine how well the model can generalize to completely unseen data and further assess the impacts of potential near-surface radar contamination from the VertiX radar.

The results of this test are shown in Fig. 4b, for each model at both sites. In general, the Z_e - S relationships show a large change in MSE (e.g., DDU MSE improves by approximately 50% at ICE-POP), while the RF MSE remains fairly consistent between sites. Furthermore, while the Z_e - S performance improves when applied at ICE-POP, the RF continues to exhibit the highest skill in predicting surface snow accumulation. While MSE values are similar for GZS and RF_{full} at GCPEX (a 5% increase in MSE for GZS), the power-law performance degrades substantially at ICE-POP with an MSE increase of 40%. We also note a similar result from IPZS where, when applied at ICE-POP (i.e., the site it was trained), IPZS demonstrates very low MSE; however, performance degrades by 150% when IPZS is applied at GCPEX. This reduction in GZS and IPZS performance when applied to new experiment sites further highlights the bespoke nature of

TABLE 3. Random forest hyperparameters and final tuned values.

Hyperparameter	Optimal value	Parameter space
Forest size	400	100, 200, 400, 800, 1600
Bootstrap	True	True, False
Max features	Sqrt	"auto," "sqrt," "log2"
Min samples split	2	1, 2, 5, 10, 20
Min samples leaf	10	1, 2, 5, 10, 20
Max depth	10	5, 10, 15, 25

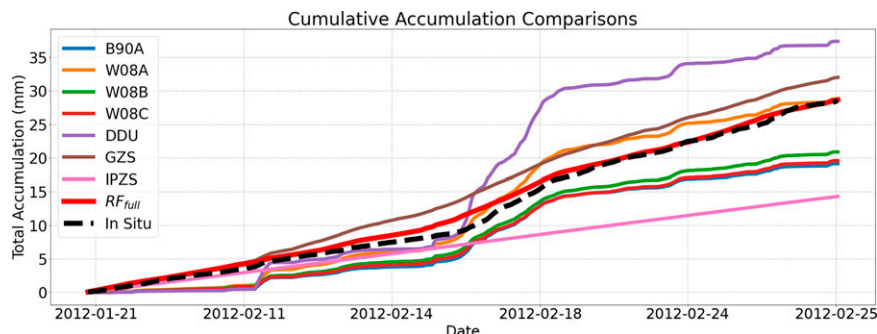


FIG. 3. Model total snow accumulation intercomparisons between in situ measurements, RF_{full} , and each respective Z_e - S relationship over the course of all accumulation events at GCPEX.

empirically derived Z_e - S relationships and their inability to generalize well to new regional climates. While the RF performance degrades slightly at ICE-POP, the overall similarity in performance suggests that the RF is capable of providing accurate estimates of accumulation, even when applied to completely unseen retrievals from a different climatic regime and time period.

While RF_{full} is able to capture the peaks in high intensity precipitation early and midway through the experiment, it struggles to accurately predict the high intensity precipitation event on 24–25 February. These high cloud storms represent a small subset of the weather events observed at GCPEX and also include much more variability in the vertical reflectivity profile observations (lower right in Fig. 2). These uncertainties in the RF estimates are further visualized in Fig. 5, which shows the distribution of model errors for the RF and each subset RF model (more details in section 3b) against the errors of the ensemble of Z_e - S relationships examined in this work. We note a slight positive bias to the RF with an anomaly curve that is centered closer to zero when compared with the negatively biased and long-tailed estimates from the Z_e - S relations in Fig. 5a (excluding GZS, which is the only positively biased Z_e - S relationship). Further, the RF model error during the

high intensity February 24–25 event is clearly visible in the error time series (Fig. 5b), where all Z_e - S relationships and the RF display high levels of error in their estimates of surface accumulation. MSE values from RF_{full} are 100% higher on average when tested on the 24–25 February storm event relative to the other storm events at GCPEX. We suggest that a larger, more robust reference dataset is required to allow the ML models to learn more clearly about differing storm types and varying cloud layer precipitation structures (which would need to be classified in advance using an additional, preprocessing supervised learning layer).

b. Vertical profile features

As a consequence of the splitting process used by individual decision tree nodes, we gain insight into the relative importance of each predictor contributing to the decisions made within the model. These predictor importance scores are useful metrics that provide additional information on model behavior and performance as a function of the atmospheric vertical column. Predictors in this study are based on the individual 221 VertiX reflectivity bins, the 221 ERA5 atmospheric temperature bins, or as combinations/aggregations of these variables. A visualization

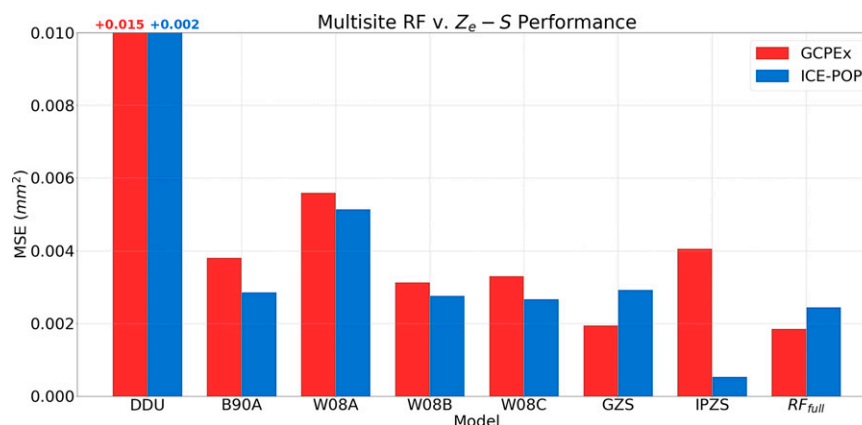


FIG. 4. MSE values from RF_{full} and the Z_e - S relationships when tested against GCPEX observations (red) and ICE-POP (blue). Note that the ICE-POP tests were trained only using GCPEX observations (excluding IPZS, which was derived from ICE-POP data) and that DDU MSE values extend beyond the chart top (extent labels are included).

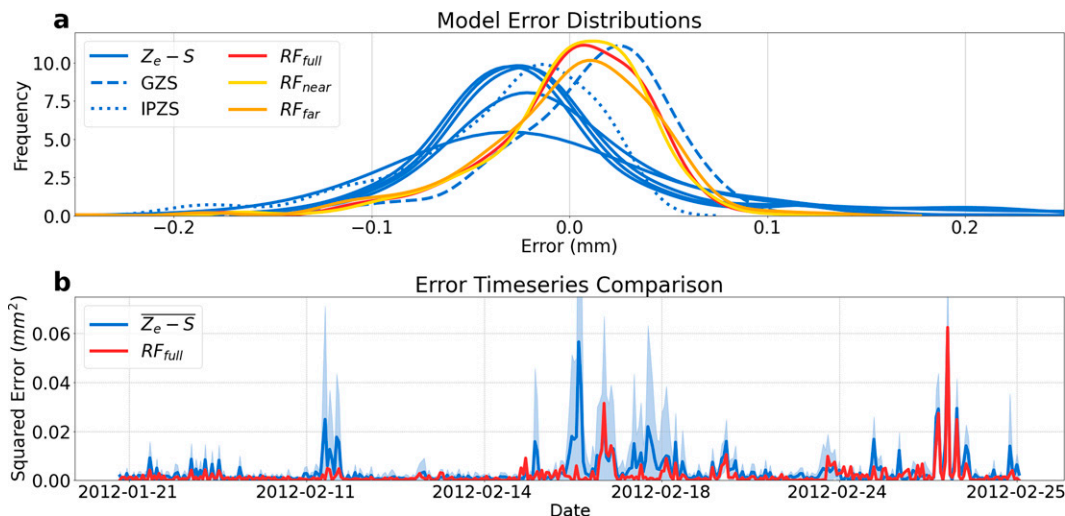


FIG. 5. (a) Distribution of model errors for RF_{full} (and subset models), along with the ensemble of $Z_e - S$ relationships for predictions at GCPEX, and (b) error time series for the RF model estimates and the mean of all tested $Z_e - S$ relationships for each 5-min time step of observed reflectivities at GCPEX (the shaded region shows a 95% confidence interval).

of the importance of each individual vertical bin for reflectivity and temperature is shown in Fig. 6, along with the respective frequency distributions for each bin in the vertical profile for RF_{full} . The frequency distribution of reflectivity values converges as we move farther up the vertical profile as missing regions are filled with the mean of the column. As noted in Fig. 6a, near-surface reflectivity values are the most significant predictors by far, followed by a handful of cloud-top level ERA5 temperature bins near 3.5 km above the surface. Reflectivity values are on average 10 times more influential in the model's decision-making process than the atmospheric temperature data.

We attempt to simplify the model architecture by refining our set of predictors from the full set of hundreds of predictors RF_{full} to two subsets (Table 4). Aggregating multiple contiguous high-

importance bins from Fig. 6 results in two new predictor sets: RF_{near} (reflectivity–temperature bins 1–45; combined importance 65%) and RF_{far} (reflectivity–temperature bins 46–221; combined importance 35%). Rerunning the model with RF_{near} as the predictor set shows very similar skill to that of RF_{full} ($MSE_{near} = 1.77 \times 10^{-3} mm^2$ vs $MSE_{full} = 1.83 \times 10^{-3} mm^2$) while using a much smaller total set of model predictors (Table 4). The RF_{far} model displays the poorest overall performance, with an approximate 19% drop in accuracy relative to using the full set of predictors.

These predictor importance scores also provide new insights into the regions of the precipitating atmospheric column that are most significant for predicting surface snow accumulation. We note in Fig. 6a, along with the results of the importance scores

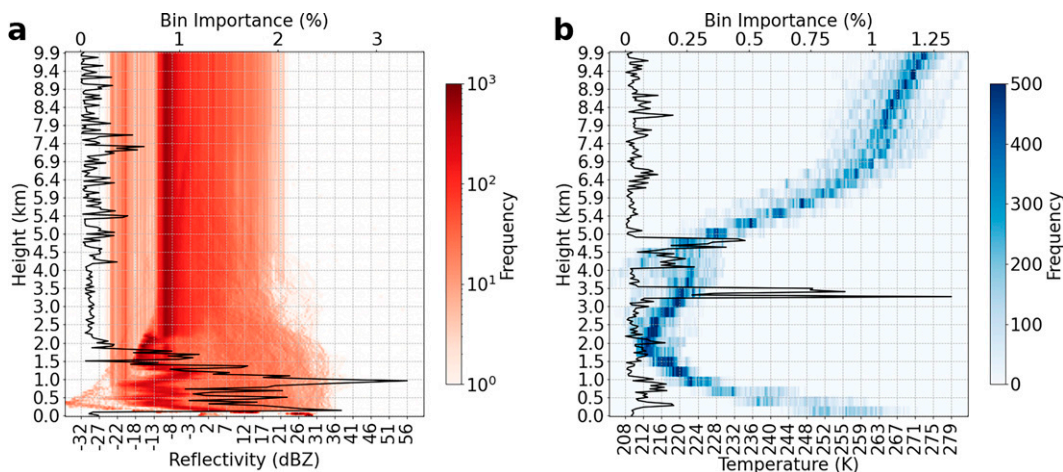


FIG. 6. Histograms of (a) reflectivity values and (b) ERA-5 temperature profiles from all snowfall events up to 10 km. The RF_{full} bin importance scores are shown in solid black for both datasets.

TABLE 4. Summary of RF models and their respective performance as a function of vertical extent.

Name	Vertical extent	MSE (mm^2)
RF _{near}	1–45	1.77×10^{-3}
RF _{far}	46–221	2.22×10^{-3}
RF _{full}	Full profile	1.83×10^{-3}

from RF_{near}, that the lowest 45 reflectivity bins (lowest 2 km of the vertical column) have a significant impact on model skill and are used in nearly two-thirds of all decisions. Meanwhile, reflectivity bins above 2 km are used much more rarely within the model, with a combined total importance of less than 22%, spanning 176 bins (~8 km) of the vertical column. These importance scores also closely correspond with the vertical profile reflectivity distributions of Fig. 6a, where there exists little activity in reflectivity values above this range and the most complex reflectivity distributions appear close to the surface. Similar importance information can be also extracted from the atmospheric temperature data; however, these feature importances are typically dwarfed by the contributions from near-surface reflectivities.

Examining distinct combinations of adjacent bins along with their relative positions within the vertical profile provides us with a better understanding of how model performance varies when trained on different subsets of the atmospheric column. Sensitivity tests varying the number of reflectivity predictors in the RF model, along with their relative position within the vertical profile, are shown in Fig. 7 for 1) MSE, 2) Pearson correlation, and 3) model bias. Each color represents a different number of adjacent predictors used in the model (i.e., red represents models trained using a sole predictor that is derived from a single bin, while dark purple represents models trained on 64 individual predictors composed of adjacent reflectivity bins in the vertical column). We note a general trend of improved performance for larger groups of adjacent predictors closer to the surface. These larger predictor-set

near-surface models have similar performance to the RF_{near} model that was generated by combining information from the most significant predictor bins.

Model performance typically degrades and eventually converges to an MSE of 2.3×10^{-3} after the 100 vertical bin range (4.5 km above the surface). Another significant feature of the sensitivity analysis is the parabolic dip in performance for bin group sizes below 32 predictors between the 40 and 90 bin ranges (2–4-km range), where we see substantial losses in model accuracy below that of even the models trained only on the uppermost bins in the atmospheric column near 10 km. While it is not entirely clear as to the cause of this performance dip in this region of the vertical profile, we believe it may be related to the fact that this region is typically between areas of shallow-cumuliform precipitation and high cloud precipitation in our training data (as seen in Fig. 2) and therefore training only on these bins provides little information as to the actual precipitation occurring within the cloud. Spaceborne radar retrievals from satellites like *CloudSat* and GPM typically are blind to the lowest 1.5 km of the atmosphere due to ground clutter interference; however, Fig. 2 shows that models trained on multiple bins within this near-surface range (bins 1–45) appear as the most significant contributors to improved predictions of surface snow accumulation. The relative performance of each of the RF models examined in this sensitivity analysis can be used to help inform current and future retrieval methods that relate a reflectivity to snowfall rate. The information encoded in multiple, adjacent near-surface bins provides a more robust model when compared with single-bin estimates or full column averages.

4. Discussion and conclusions

The VertiX radar and collocated in situ snowfall measurements from GCPEX have allowed us to derive a novel relationship between atmospheric reflectivity profiles and surface snow

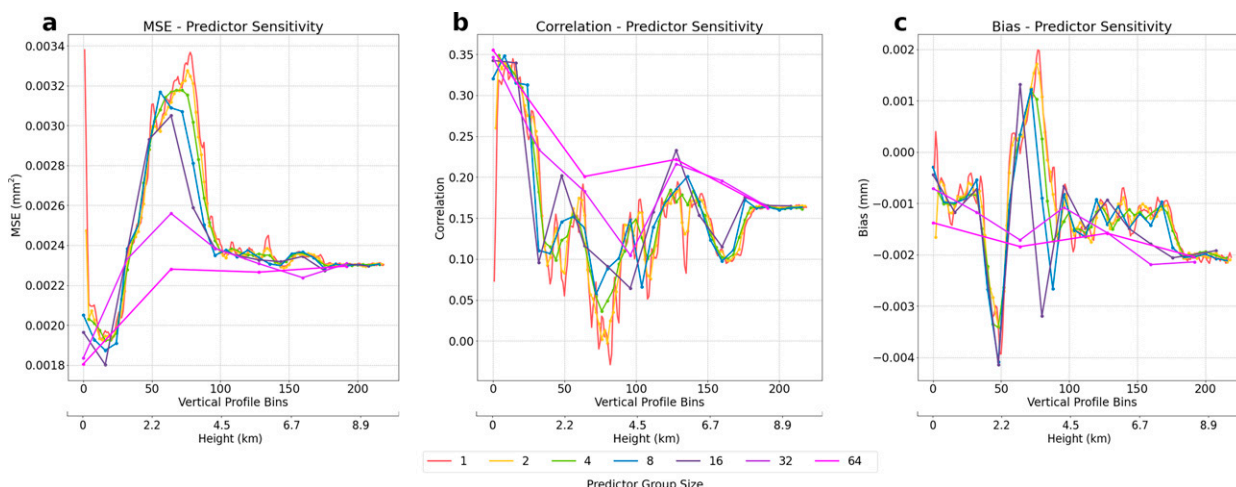


FIG. 7. Sensitivity analysis produced by varying RF predictor group sizes of adjacent bins, along with the locations of said predictor groups within the vertical profile for (a) MSE, (b) Pearson correlation, and (d) mean bias (model – observed). Each color represents a different number of predictors used to train the model at different heights in the vertical column.

accumulation using ML-based regression. The accuracy of traditional Z_e - S power-law relationships was examined when applied to the VertiX data at GCPEX, and the RF continually demonstrated the highest skill, with the Z_e - S relationships displaying clear biases in accumulation. RF predictor importance scores indicated that near-surface bins in the lowest 2 km of the atmospheric column are by far the most influential regions of the vertical profile in quantifying surface snow accumulation. Sensitivity tests comparing different combinations of RF predictors suggest that the information found within groups of multiple, adjacent near-surface bins produce a model that better understands the relationship between reflectivities and accumulation. This information is highly relevant for improving current surface snowfall retrieval algorithms from spaceborne instruments that also rely on a vertical column of reflectivity values to derive snowfall estimates for each atmospheric bin. These retrievals then extrapolate the snowfall rate from a single low-lying precipitating bin down to the surface; however, this region of the vertical column is partially masked in spaceborne radar retrievals due to ground clutter attenuation in the radar signal, contributing additional uncertainty to the true location of the lowest precipitating cloud layer.

The small sample available from GCPEX is a major limiting factor in model performance and robustness. While we have demonstrated improved performance in the ML regression models and tested the RF model against completely unseen data at ICE-POP, the highly variable nature of precipitation coupled with the effects of wind interference, the potential for mixed-phase precipitation occurrence, along with varying cloud layer precipitation structures, all contribute to increased model error and uncertainty. The vertical profile shown in Fig. 2 for event 12, for instance, depicts a variably structured, high cloud storm that is unlike the other shallow-cumuliform snowfall events being presented to the model as training data. An examination of historical in situ 2-m temperature observations also shows that surface temperatures fluctuate between $\pm 1^\circ\text{C}$ for most of this day, suggesting that mixed-phase precipitation may also have been present and contributing additional uncertainty to an already complex synoptic event. We note that these periods exhibit consistently high retrieval errors for both the RF and Z_e - S relationship (Fig. 5b), with the RF typically showing slightly reduced MSE values relative to the power-law relationships. We experimented with various methods of encoding cloud shape information and storm event types as predictors within the RF model based on similar work by Pettersen et al. (2020) to help the RF to better recognize unique snowfall event structures; however, model performance remained largely unchanged. A larger and more representative training dataset would be beneficial in teaching the RF to better identify different storm event types. Automated storm-type identification could then improve the overall model skill in predicting surface snow accumulation for a wider array of precipitation events under varying environmental and atmospheric conditions.

Although multiple layers of wind shields were employed at both GCPEX and ICE-POP, along with our attempts to examine days with only low-to-moderate wind speeds, gauge undercatch may have also been a factor contributing to error in our model because the Pluvio2 data act as the ground truth

we are attempting to predict. This is especially relevant when we consider cases of trace precipitation at 1-min temporal resolution, which have been shown in previous work by Colli et al. (2020) to be challenging to accurately quantify and correct for as a result of its low intensity. Errors in our response variable (which represents what we consider as the true accumulation occurring on the surface) would contribute to further biases in our ML regression model and further obfuscate the locations of the most important bins in the vertical profile. Uncertainties in the VertiX radar estimates should also be considered as potential contributors to model error. As previously highlighted, Fig. 2 displays a relatively uniform low-intensity near-surface reflectivity band over all days at GCPEX. It is unclear as to the origins of this feature from the VertiX retrievals, but its presence may impact the accuracy of the RF model (along with the relative importance of these bins) when predicting near-surface snowfall in this region. However, the reflectivity data from ICE-POP do not display this phenomenon and the RF continues to display similar levels of skill when tested against these unseen data, which suggests that it is not training itself on this signal or statistical noise and that its impact therefore appears to be negligible.

With these uncertainties in mind, we suggest that a follow-up study be completed that builds on the proof-of-concept methodology introduced in this work and incorporates radar data from multiple sites over an extended time period. An increased training sample would significantly contribute toward building a more robust model that could then be applied to independent, spaceborne reflectivity profiles in a wider context to derive a novel, ML-based hemispheric snow accumulation product.

Data availability statement. The GCPEX dataset is publicly available for download from NASA's Global Hydrometeorology Resource Center (GHRC; <https://ghrc.nsstc.nasa.gov/home/field-campaigns/gcpex>). ERA5 data are also publicly available for download through the European Centre for Medium-Range Weather Forecasts (ECMWF; <https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5>).

REFERENCES

- Adhikari, A., M. R. Ehsani, Y. Song, and A. Behrangi, 2020: Comparative assessment of snowfall retrieval from microwave humidity sounders using machine learning methods. *Earth Space Sci.*, **7**, e2020EA001357, <https://doi.org/10.1029/2020EA001357>.
- Berghuijs, W. R., S. Harrigan, P. Molnar, L. J. Slater, and J. W. Kirchner, 2019: The relative importance of different flood-generating mechanisms across Europe. *Water Resour. Res.*, **55**, 4582–4593, <https://doi.org/10.1029/2019WR024841>.
- Betts, A. K., D. Z. Chan, and R. L. Desjardins, 2019: Near-surface biases in ERA5 over the Canadian Prairies. *Front. Environ. Sci.*, **7**, 129, <https://doi.org/10.3389/fenvs.2019.00129>.
- Boucher, R. J., and J. G. Wieler, 1985: Radar determination of snowfall rate and accumulation. *J. Appl. Meteor. Climatol.*, **24**, 68–73, [https://doi.org/10.1175/1520-0450\(1985\)024<0068:RDOSRA>2.0.CO;2](https://doi.org/10.1175/1520-0450(1985)024<0068:RDOSRA>2.0.CO;2).

- Braham, R. R., Jr., 1990: Snow particle size spectra in lake effect snows. *J. Appl. Meteor.*, **29**, 200–207, [https://doi.org/10.1175/1520-0450\(1990\)029<0200:SPSSIL>2.0.CO;2](https://doi.org/10.1175/1520-0450(1990)029<0200:SPSSIL>2.0.CO;2).
- Buttle, J. M., and Coauthors, 2016: Flood processes in Canada: Regional and special aspects. *Canadian Water Res. J.*, **41**, 7–30, <https://doi.org/10.1080/07011784.2015.1131629>.
- Cabaj, A., P. J. Kushner, C. G. Fletcher, S. Howell, and A. A. Petty, 2020: Constraining reanalysis snowfall over the Arctic Ocean using CloudSat observations. *Geophys. Res. Lett.*, **47**, e2019GL086426, <https://doi.org/10.1029/2019GL086426>.
- Chase, R. J., S. W. Nesbitt, and G. M. McFarquhar, 2021: A dual-frequency radar retrieval of two parameters of the snowfall particle size distribution using a neural network. *J. Appl. Meteor. Climatol.*, **60**, 341–359, <https://doi.org/10.1175/JAMC-D-20-0177.1>.
- Chen, L., Y. Cao, L. Ma, and J. Zhang, 2020: A deep learning-based methodology for precipitation nowcasting with radar. *Earth Space Sci.*, **7**, e2019EA000812, <https://doi.org/10.1029/2019EA000812>.
- Colli, M., M. Stagnaro, L. G. Lanza, R. Rasmussen, and J. M. Thériault, 2020: Adjustments for wind-induced undercatch in snowfall measurements based on precipitation intensity. *J. Hydrometeorol.*, **21**, 1039–1050, <https://doi.org/10.1175/JHM-D-19-0222.1>.
- Déry, S. J., and R. D. Brown, 2007: Recent Northern Hemisphere snow cover extent trends and implications for the snow-albedo feedback. *Geophys. Res. Lett.*, **34**, L22504, <https://doi.org/10.1029/2007GL031474>.
- Duffy, G., and R. Bennartz, 2018: The role of melting snow in the ocean surface heat budget. *Geophys. Res. Lett.*, **45**, 9782–9789, <https://doi.org/10.1029/2018GL079182>.
- , F. King, R. Bennartz, and C. G. Fletcher, 2021: Seasonal estimates and uncertainties of snow accumulation from CloudSat precipitation retrievals. *Atmosphere*, **12**, 363, <https://doi.org/10.3390/atmos12030363>.
- Ehsani, M. R., A. Behrangi, A. Adhikari, Y. Song, G. J. Huffman, R. F. Adler, D. T. Bolvin, and E. J. Nelkin, 2021: Assessment of the Advanced Very High Resolution Radiometer (AVHRR) for snowfall retrieval in high latitudes using CloudSat and machine learning. *J. Hydrometeorol.*, **22**, 1591–1608, <https://doi.org/10.1175/JHM-D-20-0240.1>.
- Gatlin, P., 2020: GPM Ground Validation ICE-POP Field Campaign Data Collection. NASA Global Hydrometeorology Resource Center DAAC, accessed 1 July 2021, <https://doi.org/10.5067/GPMGV/ICEPOP/DATA101>.
- Gehring, J., A. Ferrone, A.-C. Billaulta Roux, N. Besic, K. D. Ahn, G. Lee, and A. Berne, 2021: Radar and ground-level measurements of precipitation collected by the École Polytechnique Fédérale de Lausanne during the International Collaborative Experiments for PyeongChang 2018 Olympic and Paralympic winter games. *Earth Syst. Sci. Data*, **13**, 417–433, <https://doi.org/10.5194/essd-13-417-2021>.
- Grazioli, J., C. Genthon, B. Boudevillain, C. Duran-Alarcon, M. Del Guasta, J.-B. Madeleine, and A. Berne, 2017: Measurements of precipitation in Dumont d'Urville, Adélie Land, East Antarctica. *Cryosphere*, **11**, 1797–1811, <https://doi.org/10.5194/tc-11-1797-2017>.
- Hersbach, H., and Coauthors, 2020: The ERA5 global reanalysis. *Quart. J. Roy. Meteor. Soc.*, **146**, 1999–2049, <https://doi.org/10.1002/qj.3803>.
- Hiley, M. J., M. S. Kulie, and R. Bennartz, 2011: Uncertainty analysis for CloudSat snowfall retrievals. *J. Appl. Meteor. Climatol.*, **50**, 399–418, <https://doi.org/10.1175/2010JAMC2505.1>.
- Karpatne, A., I. Ebert-Uphoff, S. Ravela, H. A. Babaie, and V. Kumar, 2019: Machine learning for the geosciences: Challenges and opportunities. *IEEE Trans. Knowl. Data Eng.*, **31**, 1544–1554, <https://doi.org/10.1109/TKDE.2018.2861006>.
- Kim, H.-U., and T.-S. Bae, 2017: Preliminary study of deep learning-based precipitation. *J. Korean Soc. Surv. Geod. Photogramm. Cartogr.*, **35**, 423–430, <https://doi.org/10.7848/ksgpc.2017.35.5.423>.
- King, F. D. M., 2019: Validating CloudSat-CPR retrievals for the estimation of snow accumulation in the Canadian Arctic. M.S. thesis, Dept. of Geography and Environmental Management, University of Waterloo, 70 pp., <https://uwspace.uwaterloo.ca/handle/10012/14736>.
- King, F., and C. G. Fletcher, 2021: Using CloudSat-derived snow accumulation estimates to constrain gridded snow water equivalent products. *Earth Space Sci.*, **8**, e2021EA001835, <https://doi.org/10.1029/2021EA001835>.
- Kochendorfer, J., and Coauthors, 2017: Analysis of single-Alter-shielded and unshielded measurements of mixed and solid precipitation from WMO-SPICE. *Hydrol. Earth Syst. Sci.*, **21**, 3525–3542, <https://doi.org/10.5194/hess-21-3525-2017>.
- , and Coauthors, 2022: How well are we measuring snow post-SPICE? *Bull. Amer. Meteor. Soc.*, **103**, E370–E388, <https://doi.org/10.1175/BAMS-D-20-0228.1>.
- Kulie, M. S., and R. Bennartz, 2009: Utilizing spaceborne radars to retrieve dry snowfall. *J. Appl. Meteor. Climatol.*, **48**, 2564–2580, <https://doi.org/10.1175/2009JAMC2193.1>.
- Lemonnier, F., and Coauthors, 2019: Evaluation of CloudSat snowfall rate profiles by a comparison with in situ micro-rain radar observations in East Antarctica. *Cryosphere*, **13**, 943–954, <https://doi.org/10.5194/tc-13-943-2019>.
- Matrosov, S. Y., K. A. Clark, and D. E. Kingsmill, 2007: A polarimetric radar approach to identify rain, melting-layer, and snow regions for applying corrections to vertical profiles of reflectivity. *J. Appl. Meteor. Climatol.*, **46**, 154–166, <https://doi.org/10.1175/JAM2508.1>.
- , C. Campbell, D. Kingsmill, and E. Sukovich, 2009: Assessing snowfall rates from X-band radar reflectivity measurements. *J. Atmos. Oceanic Technol.*, **26**, 2324–2339, <https://doi.org/10.1175/2009JTECHA1238.1>.
- McIlhattan, E. A., C. Pettersen, N. B. Wood, and T. S. L'Ecuyer, 2020: Satellite observations of snowfall regimes over the Greenland Ice Sheet. *Cryosphere*, **14**, 4379–4404, <https://doi.org/10.5194/tc-14-4379-2020>.
- Metcalfe, J. R., B. Routledge, and K. Devine, 1997: Rainfall measurement in Canada: Changing observational methods and archive adjustment procedures. *J. Climate*, **10**, 92–101, [https://doi.org/10.1175/1520-0442\(1997\)010<0092:RMICCO>2.0.CO;2](https://doi.org/10.1175/1520-0442(1997)010<0092:RMICCO>2.0.CO;2).
- Milani, L., and Coauthors, 2018: CloudSat snowfall estimates over Antarctica and the southern ocean: An assessment of independent retrieval methodologies and multi-year snowfall analysis. *Atmos. Res.*, **213**, 121–135, <https://doi.org/10.1016/j.atmosres.2018.05.015>.
- Pedregosa, F., and Coauthors, 2011: Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.*, **12**, 2825–2830.
- Peng, S., S. Piao, P. Ciais, J. Fang, and X. Wang, 2010: Change in winter snow depth and its impacts on vegetation in China. *Global Change Biol.*, **16**, 3004–3013, <https://doi.org/10.1111/j.1365-2486.2010.02210.x>.
- Pettersen, C., M. S. Kulie, L. F. Bliven, A. J. Merrelli, W. A. Petersen, T. J. Wagner, D. B. Wolff, and N. B. Wood, 2020: A composite analysis of snowfall modes from four winter seasons in Marquette, Michigan. *J. Appl. Meteor.*

- Climatol.*, **59**, 103–124, <https://doi.org/10.1175/JAMC-D-19-0099.1>.
- Rasmussen, R. M., S. Landolt, B. B. Baker, J. Kochendorfer, J. M. Thériault, and M. Colli, 2014: Examination of the performance of single alter shielded and unshielded snow gauges using observations from the Marshall Field Site during the SPICE WMO field program and numerical model simulations. *17th Symp. on Meteorological Observation and Instrumentation*, Denver, CO, Amer. Meteor. Soc., 2.4, <https://ams.confex.com/ams/21Applied17SMOI/webprogram/Paper247562.html>.
- , J. Hallett, R. Purcell, J. Cole, and M. Tryhane, 2002: The hot plate snow gauge. *11th Conf. on Cloud Physics*, Ogden, UT, Amer. Meteor. Soc., P1.6, <https://ams.confex.com/ams/11AR11CP/webprogram/Paper42751.html>.
- Schoger, S. Y., D. Moisseev, A. Lerber, S. Crewell, and K. Ebell, 2021: Snowfall-rate retrieval for K- and W-Band radar measurements designed in Hyttälä, Finland, and tested at Ny-Lesund, Svalbard, Norway. *J. Appl. Meteor. Climatol.*, **60**, 273–289, <https://doi.org/10.1175/JAMC-D-20-0095.1>.
- Sevruk, B., J. A. Hertig, and R. Spiess, 1991: The effect of a precipitation gauge orifice rim on the wind field deformation as investigated in a wind tunnel. *Atmos. Environ.*, **25A**, 1173–1179, [https://doi.org/10.1016/0960-1686\(91\)90228-Y](https://doi.org/10.1016/0960-1686(91)90228-Y).
- Shi, X., Z. Gao, L. Lausen, H. Wang, D.-Y. Yeung, W.-K. Wong, and W.-C. Woo, 2017: Deep learning for precipitation nowcasting: A benchmark and a new model. *Proc. 31st Int. Conf. on Neural Information Processing Systems*, Long Beach, CA, Neural Information Processing Systems, 5622–5632, <https://doi.org/10.48550/arXiv.1706.03458>.
- Simmons, A., and Coauthors, 2020: Global stratospheric temperature bias and other stratospheric aspects of ERA5 and ERA5.1. ECMWF Tech. Memo. 859, 40 pp.
- Skofronick-Jackson, G., and Coauthors, 2015: Global Precipitation Measurement Cold Season Precipitation Experiment (GCPEX): For measurement's sake, let it snow. *Bull. Amer. Meteor. Soc.*, **96**, 1719–1741, <https://doi.org/10.1175/BAMS-D-13-00262.1>.
- Smith, C., 2009: The relationships between snowfall catch efficiency and wind speed for the Geonor T-200B precipitation gauge utilizing various wind shield configurations. *Proc. 77th Annual Western Snow Conf.*, Canmore, AB, Canada, Western Snow Conference, 115–121.
- Stephens, G., 2017: 2C-SNOW-PROFILE-PR05. CloudSat Data Processing Center, accessed 1 July 2021, <https://www.cloudsat.cira.colostate.edu/data-products/2c-snow-profile>.
- Szyrmer, W., and I. Zawadzki, 2010: Snow studies. Part II: Average relationship between mass of snowflakes and their terminal fall velocity. *J. Atmos. Sci.*, **67**, 3319–3335, <https://doi.org/10.1175/2010JAS3390.1>.
- Tarek, M., F. P. Brissette, and R. Arsenault, 2020: Evaluation of the ERA5 reanalysis as a potential reference dataset for hydrological modelling over North America. *Hydrol. Earth Syst. Sci.*, **24**, 2527–2544, <https://doi.org/10.5194/hess-24-2527-2020>.
- Vaughan, D., and Coauthors, 2013: Observations: Cryosphere. *Climate Change 2013: The Physical Science Basis*, T. F. Stocker et al., Eds., Cambridge University Press, 317–382.
- Virtanen, P., and Coauthors, 2020: SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nat. Methods*, **17**, 261–272, <https://doi.org/10.1038/s41592-019-0686-2>.
- von Lerber, A., D. Moisseev, L. F. Bliven, W. Petersen, A.-M. Harri, and V. Chandrasekar, 2017: Microphysical properties of snow and their link to Z_e -S relations during BAEC 2014. *J. Appl. Meteor. Climatol.*, **56**, 1561–1582, <https://doi.org/10.1175/JAMC-D-16-0379.1>.
- Wang, C., R. M. Graham, K. Wang, S. Gerland, and M. A. Granskog, 2019: Comparison of ERA5 and ERA-Interim near-surface air temperature, snowfall and precipitation over Arctic sea ice: Effects on sea ice thermodynamics and evolution. *Cryosphere*, **13**, 1661–1679, <https://doi.org/10.5194/tc-13-1661-2019>.
- Wolfe, J. P., and J. R. Snider, 2012: A relationship between reflectivity and snow rate for a high-altitude S-band radar. *J. Appl. Meteor. Climatol.*, **51**, 1111–1128, <https://doi.org/10.1175/JAMC-D-11-0112.1>.
- Wood, N. B., and T. S. L'Ecuyer, 2021: What millimeter-wavelength radar reflectivity reveals about snowfall: An information-centric analysis. *Atmos. Meas. Tech.*, **14**, 869–888, <https://doi.org/10.5194/amt-14-869-2021>.
- Woods, C. P., M. T. Stoelinga, and J. D. Locatelli, 2008: Size spectra of snow particles measured in wintertime precipitation in the Pacific Northwest. *J. Atmos. Sci.*, **65**, 189–205, <https://doi.org/10.1175/2007JAS2243.1>.