



# Article Classification of Precipitation Types Based on Machine Learning Using Dual-Polarization Radar Measurements and Thermodynamic Fields

Kyuhee Shin <sup>1,2</sup>, Kwonil Kim <sup>1,2,\*</sup>, Joon Jin Song <sup>3</sup> and GyuWon Lee <sup>1,2</sup>

- <sup>1</sup> Department of Atmospheric Sciences, Kyungpook National University, Daegu 41566, Korea
- <sup>2</sup> Center for Atmospheric REmote Sensing (CARE), Kyungpook National University, Daegu 41566, Korea
- <sup>3</sup> Department of Statistical Science, Baylor University, Waco, TX 76798, USA
- Correspondence: kwonil@knu.ac.kr

**Abstract:** An accurate classification of the precipitation type is important for forecasters, particularly in the winter season. We explored the capability of three supervised machine learning (ML) methods (decision tree, random forest, and support vector machine) to determine ground precipitation types (no precipitation, rain, mixed, and snow) for winter precipitation. We provided information on the particle characteristics within a radar sampling volume and the environmental condition to the ML model with the simultaneous use of polarimetric radar variables and thermodynamic variables. The ML algorithms were optimized using predictor selection and hyperparameter tuning in order to maximize the computational efficiency and accuracy. The random forest (RF) had the highest skill scores in all precipitation types and outperformed the operational scheme. The spatial distribution of the precipitation type from the RF model showed a good agreement with the surface observation. As a result, RF is recommended for the real-time precipitation type classification due to its easy implementation, computational efficiency, and satisfactory accuracy. In addition to the validation, this study confirmed the strong dependence of precipitation type on wet-bulb temperature and a 1000–850 hPa layer thickness. The results also suggested that the base heights of the radar echo are useful in discriminating non-precipitating area.

Keywords: precipitation type; machine learning; dual-polarization radar; thermodynamic field

# 1. Introduction

Precipitation types are a major factor in weather forecasting as they affect transportation, agriculture, and industry as a whole [1,2]. Accurate discrimination among winter precipitation types (e.g., rain, snow, wet snow, and freezing rain) plays a key role in winter precipitation forecasting. An unpredicted phase transition from rain to snow, caused by small changes in the thermodynamic environment, can result in disruptions in the road and air traffic and create hazardous walking conditions. The accurate determination of surface precipitation type is also important to reduce uncertainty in hydrological models. However, the significant sensitivity of precipitation type to the variability in the atmospheric condition is one of the major challenges in the accurate classification of the type at the surface.

Several meteorological variables control the winter precipitation type. Air temperature ( $T_s$ ) is the most fundamental variable for classifying precipitation types [3,4]. In addition to the  $T_s$ , wet-bulb temperature ( $T_w$ ) is known as a more critical variable than  $T_s$  [5–8] because it accounts for the cooling effects by surface evaporation [9]. Thus,  $T_w$  generally improves the classification of the precipitation type. Relative humidity (RH) measures the amount of moisture and affects particle size distribution [5,10,11]. The thickness of the 1000–850 hPa layer (Thick<sub>1000–850</sub>) is proportional to the mean temperature of the layer. The thickness was also considered as a key variable in the discrimination between rain and snow [12–15].



Citation: Shin, K.; Kim, K.; Song, J.J.; Lee, G. Classification of Precipitation Types Based on Machine Learning Using Dual-Polarization Radar Measurements and Thermodynamic Fields. *Remote Sens.* **2022**, *14*, 3820. https://doi.org/10.3390/rs14153820

Academic Editor: Kenji Nakamura

Received: 10 June 2022 Accepted: 27 July 2022 Published: 8 August 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). The low-level vertical temperature lapse rate ( $\gamma_{low}$ ) explains the conditions in which the particles fall [6].

Based on these fundamental variables, the critical RHs as a function of  $T_s$  for the judgment of the precipitation type were obtained by statistical analysis [16] (so-called Matsuo scheme). The optimized Matsuo scheme [15], which is suitable for the Korean climate, was utilized in the Korea Meteorological Administration (KMA) operationally to support the decision making of the forecasters. In practice, the scheme is applied using the variables of the numerical weather prediction (NWP) models. Thus, the accuracy of the classification of precipitation types strongly depends on that of the NWP [17]. A low spatiotemporal resolution also causes the low predictability of the scheme.

Aside from the thermodynamical parameters, the advent of a dual-polarization radar network offers an opportunity to utilize dual-polarization variables for accurate precipitation type classification at a high spatiotemporal resolution. The dual-polarization variables include differential reflectivity ( $Z_{DR}$ , dB), specific differential phase ( $K_{DP}$ , ° km<sup>-1</sup>), and the co-polar correlation coefficient ( $\rho_{HV}$ ). In the perspective of precipitation classification,  $Z_{DR}$  and  $K_{DP}$  are useful to determine the particle shape, while  $\rho_{HV}$  is effective in determining the melting layer [18,19]. A hydrometeor classification algorithm using polarimetric variables was applied using a fuzzy-logic approach [20–22], a Bayesian approach [23,24], and so on. In particular, the fuzzy-logic algorithms were utilized for operational use and their classifications over the past 20 years are displayed in realtime in several national radar networks [25–28] to the present. Although the dual-polarization variables can provide detailed hydrometeor classification at the high spatiotemporal resolution, the absence of information (particularly thermodynamical information) below the radar beam height leads to large uncertainty in determining the precipitation type at the surface.

The synergistic use of polarimetric variables and thermodynamic variables is beneficial in the accurate classification. Schuur et al. [29] first introduced an algorithm that combines S-band polarimetric radar observations and  $T_w$  obtained from the NWP. Based on the defined Boolean decision tree, the algorithm classifies the winter precipitation phases into seven classes (crystals, dry snow, wet snow, ice pellets/sleet, freezing rain, a mix of freezing rain and ice pellets, and rain). They generated background classification from NWP and updated it using radar data. The validation with the data from an automated surface observing system (ASOS) showed an encouraging performance. Recently, Steinert et al. [30] determined the hydrometeor type at the radar beam height using a fuzzy-logic hydrometeor classification algorithm and took into account the melting process below the beam height using the NWP product based on a Boolean decision tree. Despite these efforts, the decision tree is still less objective and less flexible due to empirical critical thresholds (e.g., reflectivity,  $T_w$ , melting layer height).

Machine learning (ML) perfectly fits this particular need for modeling the complex relationship between the precipitation type and the various factors. It allows us to deal with multidimensional and complex nonlinear problems without any distributional and modeling assumptions. For these benefits, ML was widely used in the atmospheric sciences [31–33]. Using the NWP-driven variables, a multinomial logistic regression (MLR) model showed a 15% improvement in accuracy compared to the NWP forecast and outperformed the optimized Matsuo scheme [34]. Seo [35] utilized both thermodynamic variables and polarimetric variables and performed testing with the following six ML models: knearest neighbors, logistic regression, support vector machine (SVM), decision tree (DT), random forest (RF), and multi-layer perceptron. As a result, RF showed the best score and performed better than the operational method. Półrolniczak et al. [36] implemented the RF using the surface synoptic observation, NWP reanalysis data, and radar data in the similar way. However, there are opportunities to improve the performance through the optimization of each ML model. To efficiently implement the ML model in realtime operation, it is necessary to optimize and search for a suitable ML model through rigorous evaluation to ensure both accuracy and computational efficiency in realtime operation.

In this paper, we explored the capability of the optimized ML to classify the precipitation types using both polarimetric and thermodynamic variables. Three widely used supervised ML methods (DT, RF, and SVM) were tested. The optimization of the ML algorithms was performed by (1) selecting predictors through variable importance analysis and (2) tuning each ML model with the best hyperparameters. The ML models were evaluated using the hold-out validation. The best ML model was compared with the schemes in operation and applied to the Korean S-band operational radar network.

#### 2. Response and Predictors

To establish the ML model, it is necessary to properly select the response variables and predictors. In our configuration, the response variable is the precipitation type, and the predictors are both polarimetric variables and thermodynamic variables.

#### 2.1. Precipitation Type

Trained observers at KMA observed the present weather, which was recorded and assigned to one of the weather codes. The weather codes were reported at 23 stations (blue circles in Figure 1). While automatic instruments (i.e., disdrometers and present weather sensors) provided weather codes and could be used for training and validation [37], this record was deemed the most reliable source of precipitation in Korea as it was collected by a trained observer. Each observer should follow the guidelines in the observation manual of KMA to ensure the consistency of the quality from site to site.



**Figure 1.** Dual-polarization radar networks (red triangle) with ~240 km range from each radar (gray shading) and surface station for present weather observation (blue open circle).

The weather code consists of hydrometeor, lithometeor, photometeor, and electrometeor phenomena with 143 codes [38]. Among the codes, we selected 3 codes for rain (RA; rain, drizzle, and rain shower), 3 codes for snow (SN; snow, snow shower, and graupel), and 2 codes for mixed type (MIX; sleet and sleet shower). The mixed type here was defined as the co-presence of liquid (rain) and solid (snow) phases. The hail was excluded because it had never been reported in the training dataset and thus training was not possible. The other codes were considered as no precipitation (CLR), indicating the absence of precipitating particles at the surface. As a result, we divided the present weather codes reported by trained people into the following four types: rain (RA), snow (SN), mixed (MIX), and no precipitation (CLR).

# 2.2. Dual-Polarization Radar Data

The polarimetric radar variables were obtained from the radar network of 10 S-band radars [39,40] operated by the weather radar center (WRC) of the KMA. The variables include  $Z_H$ ,  $Z_{DR}$ ,  $K_{DP}$ , and  $\rho_{HV}$ . The quality of radar variables was controlled by a KMA algorithm, named the clutter elimination algorithm for the non-precipitation echo of radar data (CLEANER) [41]. The CLEANER algorithm features the efficient removal of a non-precipitation echo and allowed us to use reliable data during training.

A national 3D radar mosaic product of each radar variable has a horizontal resolution of 500 m and a vertical resolution of 50 m. To use the collocated variables as input data for ML, if any of the radar variables is missing, we considered the corresponding grid point as missing. Thus, we used data from grid points where all variables were valid. Then, we took the value of each variable at the lowest height where meteorological echo appeared for each column (i.e., the height of echo base) to reflect the characteristics of radar variables closest to the surface. We also included the height of the echo base as the predictor referred to as HGT.

Note that one of the radars (located in the most northeast in Figure 1) was replaced by a polarimetric radar from the Doppler radar in November 2019. Therefore, before November 2019,  $Z_{DR}$ ,  $K_{DP}$ , and  $\rho_{HV}$  were mosaicked by 9 radars excluding this radar, whereas  $Z_H$  was mosaicked by 10 radars.

## 2.3. Thermodynamic Field Data

The three-dimensional thermodynamic fields, which are operationally produced by KMA every 5 min, were used to retrieve thermodynamic predictors. The horizontal resolution is 4 km, and the vertical resolution is 100 m at 0–2 km height and 200 m at 2–10 km height. The fields are based on the product from the very-short-range forecast system of KMA and updated by multiquadric interpolation of observations [39,42]. The available variables include dry-bulb temperature (T), dew point temperature (T<sub>d</sub>), and pressure. The predictors were obtained from these three variables.

This study took into account the measurements at the surface and the vertical profile of the variables. In regard to the variables at the surface, we computed  $T_s$  and  $T_w$ .  $T_s$ (temperature nearest to the surface) was obtained from the T at the lowest level.  $T_w$  was calculated using  $T_s$  and RH via Stull [43], where RH was calculated from  $T_s$  and  $T_d$  at the lowest height using the "MetPy" package [44] in Python. Variables derived from the atmospheric profiles include Thick<sub>1000–850</sub> and  $\gamma_{low}$ . Thick<sub>1000–850</sub> was computed using the hypsometric equation for the layer 1000–850 hPa. We used the virtual temperature at 925 hPa for computational purposes, although the one in the equation is the layer-mean virtual temperature. Thus, Thick<sub>1000–850</sub> (gpm) is expressed as

$$\text{Thick}_{1000-850} = \frac{\text{R}_{\text{d}}\overline{\text{T}_{\text{v}}}}{g}\ln\frac{\text{P}_{1}}{\text{P}_{2}} \tag{1}$$

where  $R_d$  is the gas constant of dry air (287 J kg<sup>-1</sup> K<sup>-1</sup>), g is the gravitational acceleration (9.8 m s<sup>-2</sup>), and  $P_1$  and  $P_2$  are 1000 hPa and 850 hPa, respectively.  $\overline{T_v}$  is the virtual temperature (K) at 925 hPa.

The  $\gamma_{low}$  (K km<sup>-1</sup>) was calculated from the difference of T in the lowest 500 m height [6] as

$$\gamma_{\rm low} = -\frac{\Delta T}{\Delta z} = -\frac{T_{500m} - T_s}{0.5} \tag{2}$$

where  $T_{500m}$  is the temperature at 500 m height.

Since the horizontal resolution of dual-polarization radar measurements and thermodynamic fields are not consistent, linear interpolation was applied to the three-dimensional analysis fields prior to the computation of predictors to make the grid size the same as the dual-polarization radar data. A total of 12,367 10 min data during the winter season 2018–2019 was used for training and validation. A total of 16 days was selected including 3 days for rain-dominant, 9 days for the co-occurrence of rain, mix, and snow, and 4 days for snow-dominant cases (Table 1). In addition to the dataset for training and validation, the 5 independent cases in 2022 were selected for the application in the operational radar data.

No.	Date	Configuration	No.	Date	Configuration	
1	13 December 2018		12	15 February 2019		
2	16 December 2018		13	16 February 2019	Training	
3	23 December 2018		14	18 February 2019	and	
4	27 December 2018		15	19 February 2019	Validation	
5	28 December 2018	Training	16	27 February 2019		
6	29 December 2018	and	17	25 January 2022		
7	12 January 2019	Validation	18	15 February 2022		
8	19 January 2019		19	19 February 2022	Application	
9	31 January 2019		20	1 March 2022		
10	3 February 2019		21	19 March 2022		
11	7 February 2019		-	-	-	

Table 1. List of precipitation events for training, validation, and application.

# 3. Classification Methods of Precipitation Types

Figure 2 illustrates the overall procedures of the construction of our final ML model.



Figure 2. A flow chart for constructing the proposed machine learning model.

# 3.1. Machine Learning Methods

This study examines the effectiveness of supervised-ML methods in classifying the winter precipitation types. We considered the following three widely used classifiers: DT, RF, and SVM. The DT method is a building block of all tree-based methods. It summarizes splitting conditions used to segment the predictor region in a tree shape [45]. The DT algorithm starts by dividing the tree recursively from the top of the tree as the root node. The best predictor for splitting was selected at each node and the data were divided into

the sub-node called the intermediate node. To perform this splitting, the algorithm applies an empirical rule to decide the criteria based on the lowest impurity measure, such as the Gini index, the information gain, or the gain ratio. In this study, the Gini index was chosen for splitting nodes. The node where there is no more data to split is known as the terminal node.

RF is an ensemble model and more robust to noise in training data [46]. It consists of a number of DTs built with bootstrap samples which were generated by sampling with replacement. Each DT was grown individually using the randomly selected subset of predictors. It can reduce the variance of RF by decorrelating DTs. The final prediction was obtained by pooling predictions from all DTs. DT and RF are known as a "white box" model, which enables the interpretation of the model [47].

The SVM generally finds a hyperplane that maximizes separating the data points to their potential classes [48]. The hyperplane is defined as the boundary that divides the classes of input data. The points that are closest to the hyperplane and used to calculate the margin are called support vectors. The technique measures the distances between a new observation and each support vector with a selected kernel function and makes its prediction. In this study, we explored several kernel functions (linear, polynomial, and radial basis function) and then selected the polynomial functions for classification.

#### 3.2. Predictor Selection

In this study, a total of nine variables were regarded as the candidate predictors, which include both polarimetric ( $Z_H$ ,  $Z_{DR}$ ,  $K_{DP}$ ,  $\rho_{HV}$ , and HGT) and thermodynamic ( $T_w$ , Thick<sub>1000–850</sub>,  $T_s$ , and  $\gamma_{low}$ ) variables. Through a variable importance analysis, we examined the effect of each predictor to classify the precipitation types and selected the final predictors. The variable importance was measured by mean decrease Gini (MDG) values and was computed using the "randomForest" package [49] in R software. The MDG values are the sum of the total decrease in the Gini index from splitting on a given feature for a tree and are averaged over all trees. A large MDG value indicates an important predictor for classification.

The analysis indicated that  $T_w$  is the most important predictor for classifying the winter precipitation types (Figure 3). This is consistent with previous works that highlighted that  $T_w$  is a more significant factor in distinguishing the precipitation types than dry-bulb temperature [5,7,8], as shown by a higher MDG value (1503.05) compared to  $T_s$  (851.16). It is interesting to note that HGT is more important than the other polarimetric variables.  $\rho_{HV}$ , which is useful in hydrometeor classification, ranked third among the polarimetric variables. However,  $Z_{DR}$  and  $K_{DP}$  have the lowest importance among the candidate predictors, with MDG of 318.27 and 232.80, respectively. Thus, we excluded  $Z_{DR}$  and  $K_{DP}$  from the predictors. In other words, the final set of the predictors for the ML models includes three polarimetric ( $Z_{H}$ ,  $\rho_{HV}$ , and HGT) and four thermodynamic ( $T_w$ , Thick<sub>1000–850</sub>,  $T_s$ , and  $\gamma_{low}$ ) variables.

#### 3.3. Hyperparameter Tuning

There are several tuning parameters for each ML model. We optimized the ML models by choosing a set of optimal hyperparameters. The ML models were trained by using the selected predictors in the training and validation dataset.

DT can be optimized by reducing the depth of trees through pruning. The pruning reduces the complexity of the DT, improving the prediction accuracy by removing the overfitted subtrees. We conducted pruning using the argument called complexity parameter (cp) in the "rpart" package [50] in R software. It is based on the cost-complexity algorithm to find the optimal depth of the tree [45]. The optimal size of trees is typically chosen when cp is near 0.01. If the value is smaller than 0.01, DT tends to be complicated and overfit. The size of the tree (and corresponding cp value) is displayed in Figure 4 with the 10-fold cross-validation error (i.e., X-val relative error). The DT was appropriately pruned by choosing seven leaves to be the size of the tree.



# Variable Importance

**Figure 3.** The importance of the predictors (mean decrease Gini) with importance decreasing from top to bottom. Polarimetric variables are in green and thermodynamic variables are in gray.



**Figure 4.** X-val relative error of the DT model as a function of the different sizes of tree and corresponding cp value.

RF has two hyperparameters: the number of trees ( $n_{tree}$ ) and the number of predictors ( $m_{try}$ ). Predictors are randomly sampled in each DT. Liaw and Wiener [49] suggested  $n_{tree}$  of 500 and  $m_{try}$  of  $\sqrt{p}$  as the default value for a classification problem, where p is the total number of predictors. The default value of  $m_{try}$  is 3 ( $\approx \sqrt{7}$ ) since p is 7 in this study. To determine the optimal values of  $m_{try}$  and  $n_{tree}$ , we calculated the out-of-bag error (OOB error) by varying the  $m_{try}$  from 1 to 7 and  $n_{tree}$  from 100 to 1500 (Figure 5). The local minimum of OOB error (0.092) was found at  $m_{try}$  of 4 and  $n_{tree}$  of 500. The OOB error decreased from about 0.098 to 0.092 until  $m_{try}$  was 4 when  $n_{tree}$  of 1500 and  $m_{try}$  of 3, but time efficiency must be considered given the similar OOB error but longer computational time compared to  $n_{tree}$  of 500. Therefore, we selected  $m_{try}$  of 4 and  $n_{tree}$  of 500 as optimal values.



Figure 5. The OOB error as a function of the number of predictors  $(m_{try})$  and the number of trees  $(n_{tree})$ .

Lastly, SVM can be optimized using the cost and gamma values where the best model performance is achieved. The cost specifies the width of the margin. The cost results in a trade-off between high accuracy and smoothing decision boundary. A larger cost leads to a narrower margin (i.e., lower error) but might also increase the chance of overfitting. The gamma determines the distance of influence of a single training point. High gamma reduces the distance of influence as closer points have more weight. We evaluated the SVM model with various combinations of cost and gamma and examined the 10-fold cross-validation error rate (Figure 6). We found the optimal cost value of 0.1 and the gamma value of 3, where the lowest error rate (0.14) occurred.



**Figure 6.** The 10-fold cross-validation error rate of the SVM model under the different cost and gamma. The value in each tile indicates the error rate.

#### 3.4. Validation

The 70–30 hold-out validation was performed to compare the accuracy of ML models. The numbers of the training and validation data for each class are summarized in Table 2. A total of 8692 and 3675 10 min data were used for training and validation, respectively. It is also necessary to examine whether the ML model can provide a better classification compared to the scheme in operation. Therefore, in addition to the hold-out validation, we compared the best ML model with the optimized Matsuo scheme (see Appendix A for details), which is currently used as a fundamental reference for forecasters to make decisions [15].

Table 2. The number of 10 min data for the training and validation.

	No Precipitation (CLR)	Rain (RA)	Mixed (MIX)	Snow (SN)	Overall
Training set	3376	3721	128	1467	8692
Validation set	1425	1572	45	633	3675
Total	4801	5293	173	2100	12,367

To assess model accuracy, we used the probability of detection (POD), false alarm rate (FAR), and critical success index (CSI). The POD, FAR, and CSI were derived based on  $4 \times 4$  contingency metrics. The perfect score of POD and CSI is 1, whereas the perfect score of FAR is 0. They are defined as follows:

$$POD = \frac{H}{H+M}$$
(3)

$$FAR = \frac{F}{H+F}$$
(4)

$$CSI = \frac{H}{H + M + F}$$
(5)

where a hit (H) occurs when the predicted type agrees with the observation, a miss (M) happens when a precipitation type was observed and the model predicts another type, and a false alarm (F) represents the case when the predicted type does not occur.

# 4. Verification of Precipitation Type Classification

We evaluated the performance of ML models for each precipitation type using 70–30 holdout validation. All ML models showed a high POD for CLR, RA, and SN (Figure 7). Among the types, RA showed the best POD (0.931 for RF, 0.926 for SVM, and 0.882 for DT). This is not surprising given the huge amount of training data of RA. Despite the smaller amount of training data (less than half of RA), SN presented comparable performance to RA. In the case of MIX, all ML showed low scores (e.g., CSI less than 0.3). The smaller number of training data (1.47% of total training data) could be one of the primary reasons, but strong variability of the atmospheric condition in time and space with MIX [17] also makes the classification difficult.

For all precipitation types, the RF had the best score, followed by the SVM and the DT. The RF presented a POD of higher than 0.9 and a CSI of higher than 0.8 for both RA and SN. The RF also showed a better skill in the classification of MIX, as shown by the CSI of ~0.3, which is definitely higher than that of SVM (0.120).

As mentioned above, it is important to investigate whether ML models can perform better than the scheme in operation. The best ML model (RF) was compared with the optimized Matsuo scheme. Since the optimized Matsuo scheme (MS) does not consider no precipitation, CLR was excluded from validation.

As shown in Figure 8, the RF outperformed the optimized Matsuo scheme in the overall statistics. Both methods presented good performance in RA and SN. In particular, the PODs of the RF for RA and SN were almost equal to 1. In the MIX case, the RF showed a POD, FAR, and CSI of 0.441, 0.250, and 0.385, respectively. Although MIX is the most challenging type to predict, the RF reduced FAR by 0.742 (from 0.922 to 0.250) and increased CSI by 0.316 (from 0.069 to 0.385) compared to the optimized Matsuo scheme.



Figure 7. Skill scores of ML models, including (a) POD, (b) FAR, and (c) CSI.



**Figure 8.** Same as Figure 7, but for the optimized Matsuo scheme (MS) and RF with the precipitation events.

# 5. Application to the Operational Radar Network

We applied the best model, RF, to the Korean S-band operational radar network. The aim of this part was to present some examples of how the selected ML model can be applied to the operational radar network. The classification of precipitation types by RF over the domain is shown in Figures 9 and 10. Note that the KMA expanded the domain of the national 3D radar mosaic product and included four more S-band radars in 2022.

In the rain-dominant case at 0310 LST, 1 March 2022, a widespread precipitation system was classified into either rain or no precipitation (Figure 9a). The areas where RF predicted rain had good agreement with the ground truth. It was also noticed that the regions of no precipitation (see the zoom-in-box) are consistent with the surface observation. A good agreement with the ground truth indicated the good capability of the proposed model to identify the region of no precipitation (in addition to rain) in a widespread precipitation system.



**Figure 9.** Examples of RF classification in (**a**) rain-dominant case (0310 LST, 1 March 2022) and (**b**) snow-dominant (1420 LST, 15 February 2022). The circle filled with color indicates the ground truth. The insets in (**a**,**b**) are zoom-in-boxes, providing a clear view of the classification.



**Figure 10.** Same as Figure 9 except for cases of the co-occurrence of multiple precipitation types at (a) 0900 LST, 25 January 2022, (b) 1020 LST, 19 February 2022, and (c) 0950 LST, 19 March 2022.

The spatial distribution of the RF classification for the snow-dominant case (1420 LST, 15 February 2022) is shown in Figure 9b. Widespread snowfall dominated across the western part of the Korean Peninsula. The RF model performed well to capture the transition from no precipitation (due to sublimation) to snow. This result implies that the RF will be helpful for forecasters to determine areas of no snow but weak or moderate reflectivity.

Several cases when multiple precipitation types occurred were also examined (Figure 10). While the classification was performed at high resolution, the transition of precipitation type (from rain to snow) was consistent with the observed type (Figure 10a) even in the region where different types of precipitation were distributed in a small area.

Another example shown in Figure 10b illustrates the performance of RF in a case when no precipitation dominated. Although moderate or weak reflectivity occurred in most areas of South Korea, the RF classified a large fraction of no precipitation (Figure 10b). The result is reasonable because it coincides with the no precipitation report from the surface station. The narrow region of snow (near the bottom left region of the zoom-in-box in Figure 10b) was detected by RF and agrees with the snow report. The last example (19 March 2022) showing the result in the area of the mountainous region (Figure 10c) represented better chances of snow in the NW-SE-oriented mountain range. Besides, mixed precipitation reported at the station located near the coastal region was accurately discriminated as MIX by the RF.

Note that the trained ML model is also computationally efficient. In our case, the classification for the whole Korean radar domain was completed in a few seconds on

Intel Xeon Gold 6154 CPU at 3.00 GHz with single core, which met the computational requirement for the real-time operation. This is an encouraging result given the difficulties in performing the satisfactory classification in realtime due to the trade-off between the complexity of the algorithm (generally better accuracy) and the computational efficiency. It should be noted that the proposed model requires a thermodynamical field (e.g., variables from the NWP model) as well as polarimetric radar variables. If one datum is not available, the model is not applicable.

# 6. Discussion

Our results indicated that  $T_w$  and Thick<sub>1000-850</sub> were highly related to the precipitation types as selected in the top two variables from the variable importance analysis. This is in agreement with previous studies that recommended the use of  $T_w$  for classifying the precipitation types [5–8] since it better reflected the actual temperature of precipitation particles. This can be also seen from the distribution of  $T_w$  in the training dataset for each precipitation type (Figure 11a), which exhibited the distinctive dependence of RA, MIX, and SN on  $T_w$ . Except for CLR, the distribution of  $T_w$  of RA, MIX, and SN showed almost no overlap with each other, indicating that  $T_w$  acted as a critical factor. The narrow distribution of MIX (–1.95 to 2.64 °C) means  $T_w$  can serve as a significant factor in differentiating the most challenging type (MIX) from others.



**Figure 11.** Boxplots of (**a**) Tw, (**b**) Thick<sub>1000-850</sub>, and (**c**) HGT in terms of precipitation types using training data. A white horizontal line denotes the median, with color filled box extended from the first quartile to the third quartile.

Similarly, the distribution of Thick<sub>1000–850</sub>, which reflects the mean temperature of the layer [12,15], showed strong dependence on the precipitation type (Figure 11b). Analogous to  $T_w$ , except for CLR, Thick<sub>1000–850</sub> was also distributed in different ranges of values according to the precipitation type so that they did not overlap with each other. This result indicates that both  $T_w$  and Thick<sub>1000–850</sub> were the key factors for discriminating the precipitation types.

The presented ML model showed the reasonable identification of no precipitation areas. This could be helpful in many fields where accurate identification of the precipitation area is important such as transportation, agriculture, hydrology, and atmospheric science. To examine the evaporation or sublimation, one can utilize polarimetric variables [51–54] or/and the humidity profile from the NWP model [55]. Our study suggests that HGT can also be one of the useful factors in identifying the presence of precipitation on the ground (Figure 11c) without explicit representation of evaporation or sublimation. The HGT of CLR had higher values that are distinguishable from other types (Figure 11c). This is because the HGT implies the depth where evaporation or sublimation can occur "below the height of radar echo base".

There is still room for improvement in the model through careful treatment of noisy polarimetric variables, particularly  $Z_{DR}$  and  $K_{DP}$ . Despite the lowest importance of  $Z_{DR}$  and  $K_{DP}$  in this study, the literature suggests that these variables are still believed to be good factors to differentiate between snowfall, melting layers, and rainfall regions [56–58]. The most likely reason for the low importance is the noisiness of  $Z_{DR}$  and  $K_{DP}$ . With the reduction in the noisiness through a proper smoothing technique, the model was expected to have better performance by taking particle shape information into account.

Another thing that could lead to better performance of the model is the consideration of horizontal advection due to vertical displacement between the radar measurement and the surface, which can lead hydrometers to fall to another location on the ground. Although this study accounted for the radar variables at the lowest altitude, the location error by horizontal advection may be problematic in a strong low-level wind condition. The HGT may indirectly reflect the effect of the horizontal wind. However, one could obtain better predictions through an appropriate method to deal with this limitation.

In our study, we classified the precipitation type into CLR, RA, MIX, and SN. However, there are several important precipitation phases that we could not consider due to a lack of data such as freezing rain, ice pellets [58–62], and hail [63,64]. The ML model could be extended to include them when a sufficient number of data for these types are available.

# 7. Conclusions

Accurate classification of winter precipitation type at the surface is necessary for forecasters and researchers to minimize the damage of natural disasters on human life but is challenging due to its high variability in small changes in a thermodynamic environment. In this paper, to tackle this problem, we presented the ML-based winter precipitation type classification algorithm. The presented model attempts to exploit both polarimetric and thermodynamic variables to incorporate thermodynamical information into the classification while taking advantage of rapidly updating radar data. We tested three widely used supervised ML (DT, RF, and SVM) algorithms. The precipitation type (CLR, RA, MIX, and SN) was considered as a response. To select the important predictors, we conducted a variable importance analysis using a total of 9 candidate predictors from polarimetric radar and thermodynamical field. Subsequently, the selected 7 predictors were used in the ML models. We optimized the ML models by tuning hyperparameters and evaluated the accuracy of the three optimized models. RF was chosen as the best model.

As expected, the variable importance analysis confirmed that  $T_w$  and Thick<sub>1000-850</sub> are essential factors for distinguishing between the precipitation type (RA, MIX, and SN). This study also found the higher importance of HGT compared to other polarimetric variables. This was because HGT acted as a useful factor in discriminating no precipitation (CLR) from the precipitation type (RA, MIX, and SN).

Our fine-tuned ML models were successful in discriminating CLR, RA, MIX, and SN. The RF showed the best performance with the highest POD and the lowest FAR for all precipitation types. In particular, the RF showed the highest CSI for CLR (0.784), RA (0.846), MIX (0.283), and SN (0.840). In addition, we compared the proposed model with the optimized Matsuo scheme which is currently in operation. The results demonstrated that the ML algorithm outperformed the optimized Matsuo scheme. Despite the better performance compared to the optimized Matsuo scheme, the relatively lower score of MIX compared to other types suggests a classification of MIX remains a challenge (CSI of 0.283).

We further applied the RF to the operational S-band radar network to check the spatial distribution of the classification. The ground precipitation type was estimated at each radar grid point at a high spatial resolution (500 m). The results were generally reasonable as they showed not only the continuity in both time and space but also a good agreement with the ground truth. It should be mentioned that the ML technique is computationally efficient so the proposed model would be useful for solving realtime classification problems in other applications.

**Author Contributions:** Conceptualization, G.L. and K.S.; methodology, K.S., J.J.S. and G.L.; software, K.S. and J.J.S.; validation, K.S., J.J.S. and G.L.; formal analysis, K.S., J.J.S. and G.L.; investigation, K.S., K.K. and G.L.; resources, K.S. and K.K.; data curation, K.S. and K.K.; writing—original draft preparation, K.S.; writing—review and editing, K.K., J.J.S. and G.L.; visualization, K.S., K.K. and J.J.S.; supervision, G.L.; funding acquisition, G.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2021R1A4A1032646). This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2021R1A6A3A13042215). This work was funded by the Korea Meteorological Administration Research and Development Program under Grant KMI2022-00310.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author.

Acknowledgments: We thank the Weather Radar Center for maintaining the weather radar and for providing data. We also appreciate students and researchers in CARE, KNU for constructive discussions.

Conflicts of Interest: The authors declare no conflict of interest.

# Appendix A. The Optimized Matsuo Scheme

In this section, we presented how the optimized Matsuo scheme classifies the precipitation type. The procedure is described in [15], but we restate the methodology here since [15] is not written in English. There are two steps to determine the precipitation type: (1) discriminate snow and rain using Thick<sub>1000–850</sub> and proceed to step 2 for the undetermined types, and (2) determine the final precipitation type based on the criteria as a function of RH and T<sub>s</sub>. The criteria for each step are summarized in Table A1. Additionally, see their comprehensive figure (Figure 7 in [15]) which overlays the step 2 criteria.

**Table A1.** Criteria of the optimized Matsuo scheme to determine the precipitation type. Adapted from [15] with permission from Korean Meteorological Society.

Step	Criteria	Precipitation Type	
	$Thick_{1000-850} < 1281$	SN	
1	$1281 \le \text{Thick}_{1000-850} \le 1297$	(Proceed to step 2)	
-	$1297 < Thick_{1000-850}$	RN	
	$\label{eq:RH} \begin{array}{l} RH \geq 75 \text{ and } T_s \leq 0.9 \\ \text{and } RH < (-100/13) \times T_s + 102.5 \end{array}$	– SN – MIX	
_	$T_{s}$ > 0.9 and RH < ( $-100/13)$ $\times$ $T_{s}$ + 89.5 or $T_{s}$ $\leq$ 0.9 and RH < 75		
2	$\begin{split} RH &\geq (-100/13) \times T_s + 89.5 \\ \text{and } RH < (-100/13) \times T_s + 100 \\ \text{and } RH < (-12) \times T_s + 120 \text{ and } T_s > 0.9 \end{split}$		
_	$\begin{array}{l} T_{s} > 0.9 \text{ and } RH \geq (-100/13) \times T_{s} + 100 \\ \text{ and } RH < (-12) \times T_{s} + 120 \\ \text{ or } T_{s} \leq 0.9 \text{ and } RH \geq (-100/13) \times T_{s} + 102.5 \end{array}$		
	$RH \geq (-12) \times T_s + 120$	RN	

# References

- Norrman, J.; Eriksson, M.; Lindqvist, S. Relationships between Road Slipperiness, Traffic Accident Risk and Winter Road Maintenance Activity. *Clim. Res.* 2000, 15, 185–193. [CrossRef]
- Harpold, A.A.; Kaplan, M.L.; Klos, P.Z.; Link, T.; McNamara, J.P.; Rajagopal, S.; Schumer, R.; Steele, C.M. Rain or Snow: Hydrologic Processes, Observations, Prediction, and Research Needs. *Hydrol. Earth Syst. Sci.* 2017, 21, 1–22. [CrossRef]
- Habets, F.; Boone, A.; Champeaux, J.L.; Etchevers, P.; Franchistéguy, L.; Leblois, E.; Ledoux, E.; Le Moigne, P.; Martin, E.; Morel, S.; et al. The SAFRAN-ISBA-MODCOU Hydrometeorological Model Applied over France. J. Geophys. Res. Atmos. 2008, 113, D06113. [CrossRef]

- 4. Zhong, K.; Zheng, F.; Xu, X.; Qin, C. Discriminating the Precipitation Phase Based on Different Temperature Thresholds in the Songhua River Basin, China. *Atmos. Res.* **2018**, 205, 48–59. [CrossRef]
- 5. Ding, B.; Yang, K.; Qin, J.; Wang, L.; Chen, Y.; He, X. The Dependence of Precipitation Types on Surface Elevation and Meteorological Conditions and Its Parameterization. *J. Hydrol.* **2014**, *513*, 154–163. [CrossRef]
- 6. Sims, E.M.; Liu, G. A Parameterization of the Probability of Snow–Rain Transition. *J. Hydrometeorol.* **2015**, *16*, 1466–1477. [CrossRef]
- Behrangi, A.; Yin, X.; Rajagopal, S.; Stampoulis, D.; Ye, H. On Distinguishing Snowfall from Rainfall Using Near-Surface Atmospheric Information: Comparative Analysis, Uncertainties and Hydrologic Importance. *Q. J. R. Meteorol. Soc.* 2018, 144, 89–102. [CrossRef]
- Wang, Y.H.; Broxton, P.; Fang, Y.; Behrangi, A.; Barlage, M.; Zeng, X.; Niu, G.Y. A Wet-Bulb Temperature-Based Rain-Snow Partitioning Scheme Improves Snowpack Prediction Over the Drier Western United States. *Geophys. Res. Lett.* 2019, 46, 13825–13835.
  [CrossRef]
- 9. Rogers, R.R.; Yau, M.K. A Short Course in Cloud Physics, 3rd ed.; Elsevier: Amsterdam, The Netherlands, 1996; ISBN 0-08-057094-1.
- 10. Thériault, J.M.; Stewart, R.E.; Henson, W. On the Dependence of Winter Precipitation Types on Temperature, Precipitation Rate, and Associated Features. *J. Appl. Meteorol. Clim.* **2010**, *49*, 1429–1442. [CrossRef]
- 11. Froidurot, S.; Zin, I.; Hingray, B.; Gautheron, A. Sensitivity of Precipitation Phase over the Swiss Alps to Different Meteorological Variables. *J. Hydrometeorol.* **2014**, *15*, 685–696. [CrossRef]
- 12. Keeter, K.K.; Cline, J.W. The Objective Use of Observed and Forecast Thickness Values to Predict Precipitation Type in North Carolina. *Weather Forecast.* **1991**, *6*, 456–469. [CrossRef]
- 13. Heppner, P.O.G. Snow versus Rain: Looking beyond the "Magic" Numbers. Weather Forecast. 1992, 7, 683–691. [CrossRef]
- 14. Bourgouin, P. A Method to Determine Precipitation Types. Weather Forecast. 2000, 15, 583–592. [CrossRef]
- 15. Lee, S.-M.; Han, S.-U.; Won, H.Y.; Ha, J.-C.; Lee, Y.H.; Lee, J.-H.; Park, J.-C. A Method for the Discrimination of Precipitation Type Using Thickness and Improved Matsuo's Scheme over South Korea. *Atmosphere* **2014**, *24*, 151–158. [CrossRef]
- Matsuo, T.; Sasyo, Y.; Sato, Y. Relationship between Types of Precipitation on the Ground and Surface Meteorological Elements. J. Meteorol. Soc. Jpn. Ser. II 1981, 59, 462–476. [CrossRef]
- 17. Reeves, H.D.; Elmore, K.L.; Ryzhkov, A.; Schuur, T.; Krause, J. Sources of Uncertainty in Precipitation-Type Forecasting. *Weather Forecast.* 2014, 29, 936–953. [CrossRef]
- 18. Bringi, V.N.; Chandrasekar, V. *Polarimetric Doppler Weather Radar: Principles and Applications*; Cambridge University Press: Cambridge, UK, 2001; ISBN 978-0-521-62384-1.
- 19. Kumjian, M.R. Principles and Applications of Dual-Polarization Weather Radar. Part II: Warm and Cold Season Applications. *J. Oper. Meteorol.* **2013**, *1*, 243–264. [CrossRef]
- 20. Park, H.S.; Ryzhkov, A.V.; Zrnić, D.S.; Kim, K.E. The Hydrometeor Classification Algorithm for the Polarimetric WSR-88D: Description and Application to an MCS. *Weather Forecast.* **2009**, *24*, 730–748. [CrossRef]
- 21. Al-Sakka, H.; Boumahmoud, A.-A.; Fradon, B.; Frasier, S.J.; Tabary, P. A New Fuzzy Logic Hydrometeor Classification Scheme Applied to the French X-, C-, and S-Band Polarimetric Radars. *J. Appl. Meteorol. Clim.* **2013**, *52*, 2328–2344. [CrossRef]
- 22. Dolan, B.; Rutledge, S.A.; Lim, S.; Chandrasekar, V.; Thurai, M. A Robust C-Band Hydrometeor Identification Algorithm and Application to a Long-Term Polarimetric Radar Dataset. *J. Appl. Meteorol. Clim.* **2013**, *52*, 2162–2186. [CrossRef]
- Marzano, F.S.; Scaranari, D.; Montopoli, M.; Vulpiani, G. Supervised Classification and Estimation of Hydrometeors from C-Band Dual-Polarized Radars: A Bayesian Approach. *IEEE Trans. Geosci. Remote Sens.* 2008, 46, 85–98. [CrossRef]
- 24. Yang, J.; Zhao, K.; Zhang, G.; Chen, G.; Huang, H.; Chen, H. A Bayesian Hydrometeor Classification Algorithm for C-Band Polarimetric Radar. *Remote Sens.* **2019**, *11*, 1884. [CrossRef]
- Vivekanandan, J.; Ellis, S.M.; Oye, R.; Zrnic, D.S.; Ryzhkov, A.V.; Straka, J. Cloud Microphysics Retrieval Using S-Band Dual-Polarization Radar Measurements. *Bull. Am. Meteor. Soc.* 1999, 80, 381–388. [CrossRef]
- Liu, H.; Chandrasekar, V. Classification of Hydrometeors Based on Polarimetric Radar Measurements: Development of Fuzzy Logic and Neuro-Fuzzy Systems, and In Situ Verification. J. Atmos. Ocean. Technol. 2000, 17, 140–164. [CrossRef]
- 27. Straka, J.M.; Zrnić, D.S.; Ryzhkov, A.V. Bulk Hydrometeor Classification and Quantification Using Polarimetric Radar Data: Synthesis of Relations. *J. Appl. Meteor.* **2000**, *39*, 1341–1372. [CrossRef]
- Rico-Ramirez, M.A.; Cluckie, I.D.; Han, D. Correction of the Bright Band Using Dual-Polarisation Radar. Atmos. Sci. Lett. 2005, 6, 40–46. [CrossRef]
- Schuur, T.J.; Park, H.-S.; Ryzhkov, A.V.; Reeves, H.D. Classification of Precipitation Types during Transitional Winter Weather Using the RUC Model and Polarimetric Radar Retrievals. J. Appl. Meteorol. Clim. 2012, 51, 763–779. [CrossRef]
- 30. Steinert, J.; Tracksdorf, P.; Heizenreder, D. Hymec: Surface Precipitation Type Estimation at the German Weather Service. *Weather Forecast.* **2021**, *36*, 1611–1627. [CrossRef]
- McGovern, A.; Elmore, K.L.; Gagne, D.J.; Haupt, S.E.; Karstens, C.D.; Lagerquist, R.; Smith, T.; Williams, J.K. Using Artificial Intelligence to Improve Real-Time Decision-Making for High-Impact Weather. *Bull. Am. Meteorol. Soc.* 2017, 98, 2073–2090. [CrossRef]
- 32. Jergensen, G.E.; McGovern, A.; Lagerquist, R.; Smith, T. Classifying Convective Storms Using Machine Learning. *Weather Forecast*. 2020, *35*, 537–559. [CrossRef]

- 33. Shin, K.; Song, J.J.; Bang, W.; Lee, G. Quantitative Precipitation Estimates Using Machine Learning Approaches with Operational Dual-Polarization Radar Data. *Remote Sens.* **2021**, *13*, 694. [CrossRef]
- Moon, S.H.; Kim, Y.H. An Improved Forecast of Precipitation Type Using Correlation-Based Feature Selection and Multinomial Logistic Regression. *Atmos. Res.* 2020, 240, 104928. [CrossRef]
- Seo, B.C. A Data-Driven Approach for Winter Precipitation Classification Using Weather Radar and NWP Data. Atmosphere 2020, 11, 701. [CrossRef]
- Półrolniczak, M.; Kolendowicz, L.; Czernecki, B.; Taszarek, M.; Tóth, G. Determination of Surface Precipitation Type Based on the Data Fusion Approach. *Adv. Atmos. Sci.* 2021, *38*, 387–399. [CrossRef]
- 37. Pickering, B.S.; Best, S.; Dufton, D.; Lukach, M.; Lyth, D.; Neely, R.R. Improving Observations of Precipitation Type at the Surface: A 5-Year Verification of a Radar-Derived Product from the United Kingdom's Met Office. *J. Hydrometeorol.* **2021**, 22, 19. [CrossRef]
- Korea Meteorological Administration Manual of Surface Weather Observation. 2016. Available online: https://book.kma.go.kr/ viewer/MediaViewer.ax?cid=33393&rid=5&moi=5241 (accessed on 28 July 2022).
- 39. Lee, J.E.; Jung, S.H.; Kwon, S. Characteristics of the Bright Band Based on Quasi-Vertical Profiles of Polarimetric Observations from an s-Band Weather Radar Network. *Remote Sens.* **2020**, *12*, 4061. [CrossRef]
- 40. Lee, J.-E.; Kwon, S.; Jung, S.-H. Real-Time Calibration and Monitoring of Radar Reflectivity on Nationwide Dual-Polarization Weather Radar Network. *Remote Sens.* **2021**, *13*, 2936. [CrossRef]
- 41. Oh, Y.A.; Kim, H.L.; Suk, M.K. Clutter Elimination Algorithm for Non-Precipitation Echo of Radar Data Considering Meteorological and Observational Properties in Polarimetric Measurements. *Remote Sens.* **2020**, *12*, 3790. [CrossRef]
- 42. Kim, M.; Lee, K.; Lee, Y.H. Visibility Data Assimilation and Prediction Using an Observation Network in South Korea. *Pure Appl. Geophys.* 2020, 177, 1125–1141. [CrossRef]
- 43. Stull, R. Wet-Bulb Temperature from Relative Humidity and Air Temperature. J. Appl. Meteorol. Clim. 2011, 50, 2267–2269. [CrossRef]
- May, R.M.; Arms, S.C.; Marsh, P.; Bruning, E.; Leeman, J.R.; Goebbert, K.; Thielen, J.E.; Bruick, Z.S.; Camron, M.D. MetPy: A Python Package for Meteorological Data. Unidata 2022. [CrossRef]
- 45. Breiman, L.; Friedman, J.; Stone, C.J.; Olshen, R.A. Classification and Regression Trees; Routledge: New York, NY, USA, 1984.
- 46. Breiman, L. Random Forests. Mach. Learn. 2001, 45, 5–32. [CrossRef]
- McGovern, A.; Lagerquist, R.; Gagne, D.J.; Jergensen, G.E.; Elmore, K.L.; Homeyer, C.R.; Smith, T. Making the Black Box More Transparent: Understanding the Physical Implications of Machine Learning. *Bull. Am. Meteorol. Soc.* 2019, 100, 2175–2199. [CrossRef]
- 48. Cortes, C.; Vapnik, V. Support Vector Networks. *Mach. Learn.* **1995**, 20, 273–297. [CrossRef]
- 49. Liaw, A.; Wiener, M. Classification and Regression by Random Forest. R News 2002, 2, 18–22.
- 50. Therneau, T.; Atkinson, B.; Ripley, B. Rpart: Recursive Partitioning and Regression Trees. 2010. Available online: https://cran.r-project.org/web/packages/rpart/ (accessed on 28 July 2022).
- 51. Kumjian, M.R.; Ryzhkov, A.V. The Impact of Evaporation on Polarimetric Characteristics of Rain: Theoretical Model and Practical Implications. *J. Appl. Meteorol. Clim.* **2010**, *49*, 1247–1267. [CrossRef]
- Xie, X.; Evaristo, R.; Troemel, S.; Saavedra, P.; Simmer, C.; Ryzhkov, A. Radar Observation of Evaporation and Implications for Quantitative Precipitation and Cooling Rate Estimation. J. Atmos. Ocean. Technol. 2016, 33, 1779–1792. [CrossRef]
- Carlin, J.T.; Reeves, H.D.; Ryzhkov, A.V. Polarimetric Observations and Simulations of Sublimating Snow: Implications for Nowcasting. J. Appl. Meteorol. Clim. 2021, 60, 1035–1054. [CrossRef]
- 54. Song, J.J.; Innerst, M.; Shin, K.; Ye, B.; Kim, M.; Yeom, D.; Lee, G. Estimation of Precipitation Area Using S-Band Dual-Polarization Radar Measurements. *Remote Sens.* **2021**, *13*, 2039. [CrossRef]
- 55. Martinaitis, S.M.; Grams, H.M.; Langston, C.; Zhang, J.; Howard, K. A Real-Time Evaporation Correction Scheme for Radar-Derived Mosaicked Precipitation Estimations. *J. Hydrometeorol.* **2018**, *19*, 87–111. [CrossRef]
- Ryzhkov, A.V.; Kumjian, M.R.; Ganson, S.M.; Khain, A.P. Polarimetric Radar Characteristics of Melting Hail. Part I: Theoretical Simulations Using Spectral Microphysical Modeling. J. Appl. Meteorol. Clim. 2013, 52, 2849–2870. [CrossRef]
- Trömel, S.; Ryzhkov, A.V.; Zhang, P.; Simmer, C. Investigations of Backscatter Differential Phase in the Melting Layer. J. Appl. Meteorol. Clim. 2014, 53, 2344–2359. [CrossRef]
- Tobin, D.M.; Kumjian, M.R. Polarimetric Radar and Surface-Based Precipitation-Type Observations of Ice Pellet to Freezing Rain Transitions. Weather Forecast. 2017, 32, 2065–2082. [CrossRef]
- Reeves, H.D.; Ryzhkov, A.V.; Krause, J. Discrimination between Winter Precipitation Types Based on Spectral-Bin Microphysical Modeling. J. Appl. Meteorol. Clim. 2016, 55, 1747–1761. [CrossRef]
- 60. McCray, C.D.; Atallah, E.H.; Gyakum, J.R. Long-Duration Freezing Rain Events over North America: Regional Climatology and Thermodynamic Evolution. *Weather Forecast.* **2019**, *34*, 665–681. [CrossRef]
- 61. Tripp, D.D.; Martin, E.R.; Reeves, H.D. Applications of Uncrewed Aerial Vehicles (UAVs) in Winter Precipitation-Type Forecasts. J. Appl. Meteorol. Clim. 2021, 60, 361–375. [CrossRef]
- 62. Lachapelle, M.; Thériault, J.M. Characteristics of Precipitation Particles and Microphysical Processes during the 11–12 January 2020 Ice Pellet Storm in the Montréal Area, Québec, Canada. *Mon. Weather Rev.* **2022**, *150*, 1043–1059. [CrossRef]

- 63. Punge, H.J.; Kunz, M. Hail Observations and Hailstorm Characteristics in Europe: A Review. *Atmos. Res.* 2016, 176–177, 159–184. [CrossRef]
- 64. Allen, J.T.; Giammanco, I.M.; Kumjian, M.R.; Punge, H.J.; Zhang, Q.; Groenemeijer, P.; Kunz, M.; Ortega, K. Understanding Hail in the Earth System. *Rev. Geophys.* 2020, *58*, e2019RG000665. [CrossRef]