

Article

Precipitation Phase Classification with X-Band Polarimetric Radar and Machine Learning Using Micro Rain Radar and Disdrometer Data in Grenoble (French Alps)

Francesc Polls ¹ , Brice Boudevillain ² , Mireia Udina ¹ , Francisco J. Ruiz ³ , Albert Garcia-Benadi ⁴ , Eulàlia Busquets ¹ , Matthieu Vernay ⁵ and Joan Bech ^{1,6,*} 

¹ Department Applied Physics—Meteorology, Universitat de Barcelona, Martí Franquès 1, 08028 Barcelona, Spain; fpolls@meteo.ub.edu (F.P.); mudina@meteo.ub.edu (M.U.); ebusquets@meteo.ub.edu (E.B.)

² Institute for Geosciences and Environmental Research, Université Grenoble Alpes, CNRS, INRAE, IRD, Grenoble INP, IGE, 38000 Grenoble, France; brice.boudevillain@univ-grenoble-alpes.fr

³ Automatic Control, University: UPC-BarcelonaTech, Polytechnic School of Engineering of Vilanova i la Geltrú, Av. Víctor Balaguer 1, 08800 Vilanova i la Geltrú, Spain; francisco.javier.ruiz@upc.edu

⁴ Physics Department, EPSEVEG, UPC, 08800 Vilanova i la Geltrú, Spain; albert.garcia-benadi@upc.edu

⁵ Centre d'Études de la Neige, Université de Toulouse, Université Grenoble Alpes, Météo-France, CNRS, CNRM, 38400 Grenoble, France; matthieu.vernay@meteo.fr

⁶ Water Research Institute, Universitat de Barcelona, 08028 Barcelona, Spain

* Correspondence: joan.bech@ub.edu

Highlights

What are the main findings?

- Existing precipitation phase classification schemes (solid, liquid and mixed) based on Micro Rain Radar (MRR) and X-band weather radar polarimetric observations are compared with disdrometer data and the former achieves best results (86% vs. 47% accuracy).
- Based on comparison of previous schemes, a new methodology is developed for X-Band polarimetric radar using machine learning random forest trained with MRR classifications, achieving an accuracy of 78%.

What are the implications of the main findings?

- The new precipitation phase classification method represents a significant improvement in accuracy compared to previous methods for polarimetric radar data.
- The approach offers valuable insights for operational forecasting, water resource management, and other applications in mountainous regions or areas where precipitation phase transitions need to be monitored.



Academic Editors: Dong-In Lee and Kaoshen Chung

Received: 15 December 2025

Revised: 17 January 2026

Accepted: 23 January 2026

Published: 29 January 2026

Copyright: © 2026 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 \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Abstract

Accurate classification of precipitation phase (liquid, mixed, or solid) is essential in high mountain environments, where rapid changes in elevation can lead to abrupt phase transitions over short distances, significantly affecting hydro-meteorological, ecological, and socio-economic activities. However, most existing classification schemes have not been evaluated over long periods using real observational data, but mainly through simulations. This study addresses this gap by introducing a new methodology based on X-band polarimetric radar and by validating it against real precipitation events over an extended time period. The machine learning model is trained and tested using a four-year dataset including X-band radar, Micro Rain Radar, disdrometer, and temperature profile data from the Grenoble region (French Alps). To improve the classification accuracy, three temperature profile

sources were tested: lapse rates obtained from automatic weather stations, interpolation of the temperature profile from the freezing level detected by the Micro Rain Radar, and temperature profiles from the operational AROME model forecast. Three different phase classification schemes were tested: two existing schemes based on fuzzy-logic, and the new method based on random forest. Results show that the random forest method, trained with radar polarimetric variables, AROME temperature profiles, and target labels derived from Micro Rain Radar observations, achieves the highest accuracy. Despite the overall good classification results, limitations persist in identifying mixed-phase precipitation due to its transitional nature and vertical variability. Feature importance analysis indicates that temperature is the most influential variable in the classification scheme, followed by reflectivity factor measured in the horizontal plane (Z_e) and differential reflectivity (Z_{dr}). This methodology demonstrates the potential of combining machine learning techniques with multi-instrument observations to improve hydrometeor classification in complex terrain. The approach offers valuable insights for operational forecasting, water resource management, and climate impact assessments in mountainous regions.

Keywords: X-band; dual-polarization; micro rain radar; disdrometer; machine learning; random forest; precipitation type; Doppler radar

1. Introduction

Accurate identification of precipitation phase (liquid, solid, or mixed) is crucial in high mountain environments, influencing drinking water supply and hydropower generation to agriculture and winter tourism [1,2]. Orography strongly influences air flow patterns, altering the spatial and temporal distribution of precipitation. Specifically, rapid changes in terrain elevation over short distances play a key role in determining the spatial variability of precipitation phase [3]. In many mountain regions, these contrasts affect snowpack evolution, water availability, and the timing of river flows, impacts that have led to shifts such as earlier streamflow timing—by about one month in recent decades [4]—forcing ski resorts, a key economic driver in many mountain areas, to increasingly rely on artificial snow production [5]. The Alps provide a clear example of this complexity: sharp slopes and rapid temperature changes associated with cold and warm fronts often lead to abrupt transitions between liquid and solid precipitation within just a few meters of elevation [6,7].

However, having an automated method to differentiate among the various precipitation types remains challenging, particularly when the goal is to classify phase over a broad area rather than at a single observation point. This spatial extension is critical for understanding how precipitation evolves across mountainous regions, yet it is limited by the capabilities of both ground-based and remote sensing instruments, especially under transitional or mixed-phase precipitation conditions [8,9].

Instruments such as disdrometers provide continuous in situ measurements of precipitation phase by directly analyzing precipitation particle characteristics. However, these instruments only offer point-scale observations. To extend phase classification over larger areas, operational polarimetric weather radars have been widely employed. These radar-based phase estimates rely on indirect approaches, e.g., fuzzy logic methodologies developed from scattering simulations [10,11]. These techniques use membership functions defined for each hydrometeor type and each polarimetric variable, exploiting characteristic patterns that vary with precipitation type within the sampled volume. Other approaches combine fuzzy logic schemes with unsupervised statistical clustering methods [12]. Recently, other studies have started to apply machine learning techniques

to address hydrometeor type classification. For example, [13] proposes a data-driven approach for winter precipitation classification using weather radar and numerical weather prediction data, whereas [14] develops a methodology based on dual-polarization radar measurements and thermodynamic fields. Both studies focus on surface hydrometeor classification using temperature information and polarimetric radar data.

A major challenge in radar-based phase classification is the evaluation of these methods. The lack of in situ measurements above the surface severely limits validation of above ground-level data. Alternatively, some studies have employed vertically pointing Micro Rain Radar (MRR) observations for hydrometeor classification purposes with promising results [15–17]. However, these approaches have only been tested with co-located ground-level disdrometer observations, not at higher altitudes.

For this reason, the objective of this study is three-fold: first, to test existing hydrometeor classification methodologies for MRR using disdrometer data at different altitudes. Second, to test X-band weather radar-based classification methods using both disdrometer and MRR observations. Initial tests of the X-band classification methodology showed good results, but with room for improvement. This motivated the third objective: to develop a novel classification model to improve upon previous approaches by using a supervised machine learning technique based on a random forest algorithm trained with MRR classifications. MRR was chosen as the ground truth reference because it provides real-time, high-resolution measurements of precipitation profiles. Alternative sources such as reanalysis or model outputs have much coarser resolution, and radiosonde observations are both unavailable in the study area and temporally sparse. Moreover, real-time applicability is a key objective of this work, which requires instruments capable of delivering immediate data rather than post-processed products.

2. Materials and Methods

2.1. Region of Study

The Grenoble region, located in the French Alps, lies within a Y-shaped alluvial valley at an average elevation of approximately 220 m above sea level (ASL) [7]. It is enclosed by three major mountain ranges: the Chartreuse to the north, the Belledonne to the southeast, and the Vercors to the west (reaching 2083, 2977, and 2341 m ASL, respectively). This complex topography forms a deep basin that strongly influences the local meteorological conditions [18]. The dominant winds are strongly influenced by the valley's channelling, leading to significant variations depending on the location [19]. The region experiences a semi-continental climate, classified as Cfb under the Köppen system [20] and characterized by moderate precipitation distributed throughout the year, relatively cold winters, and warm summers. Due to the basin effect caused by the enclosing mountain ranges, the Grenoble valley tends to exhibit warmer conditions than surrounding elevated areas, resulting in a distinct warm-temperate microclimate. Annual precipitation in the region typically averages around 1000 mm in the valley bottom and can be double that amount in the surrounding mountains [18].

Figure 1 shows a topographic map of the study area, including the location of the Météo-France X-band radar system (hereafter XRAD) situated at the summit of Mount Moucherotte in the Vercors range, and the Institute of Environmental Geosciences research center (IGE) located on the valley floor. At the IGE site, continuous in situ observations are conducted using an MRR, a disdrometer, and an automatic weather station (AWS). Figure 1 also includes the Saint-Martin-d'Uriage site (904 m ASL) and the Chamrousse site (1730 m ASL), both equipped with an IGE disdrometer. The Chamrousse site additionally hosts a Météo-France AWS. Ten other Météo-France AWS stations are distributed along the valley within 25 km of the IGE site (three of them shown in Figure 1).

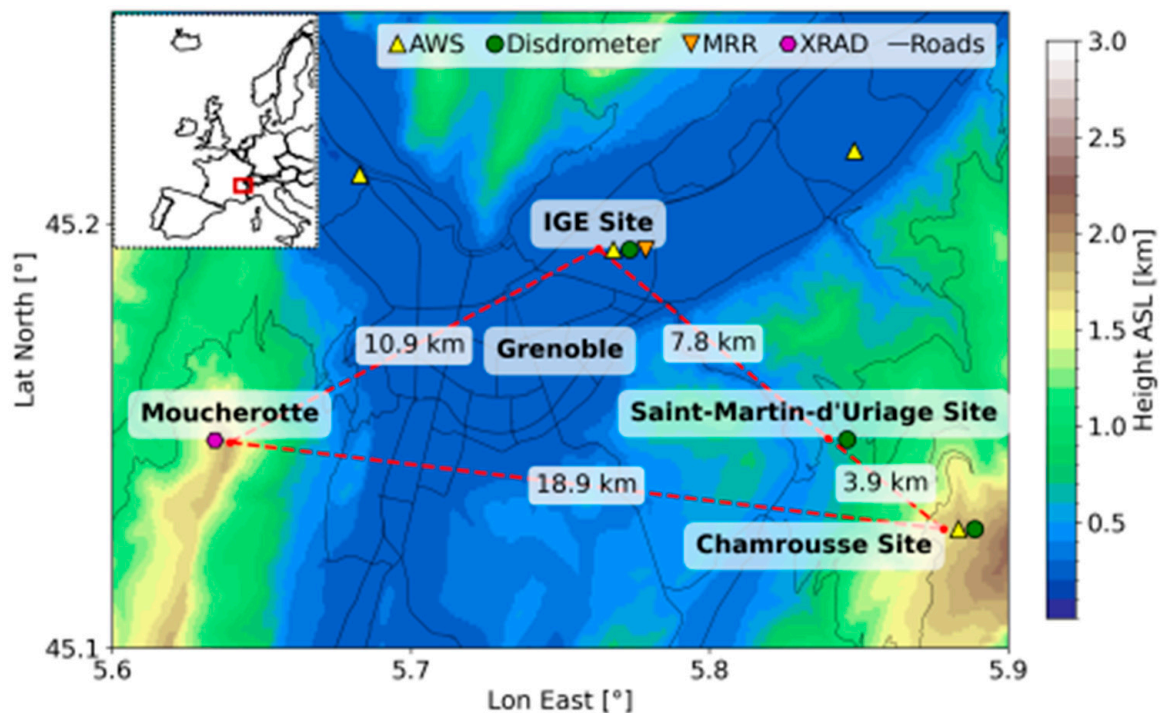


Figure 1. Terrain elevation of the study region (in km above sea level, color-shaded). The map shows main roads and instrumented sites including Mount Moucherotte (1901 m ASL, X-band radar), the IGE center (220 m ASL, disdrometer, MRR, and AWS), Chamrousse (1730 m ASL, disdrometer and AWS), and Saint-Martin-d'Uriage (904 m ASL, disdrometer). Distances between sites (dashed lines) are also shown. Top left, western Europe map with the study region marked with a red rectangle.

Figure 1 also indicates that horizontal distances between the study sites are less than 20 km. Given the terrain geometry, there is no radar beam blockage from the XRAD site, which is crucial for the analysis performed. This configuration ensures that the MRR at the IGE site and the disdrometers at Chamrousse and Saint-Martin-d'Uriage can be reliably compared on days when precipitation is sufficiently widespread across the area. The absence of radar beam blockage of the XRAD radar over the IGE and Chamrousse sites is detailed in Figure 2, which includes two vertical cross-sections and a schematic view of the XRAD radar's Plan Position Indicator (PPI) scans relative to the MRR antenna beam.

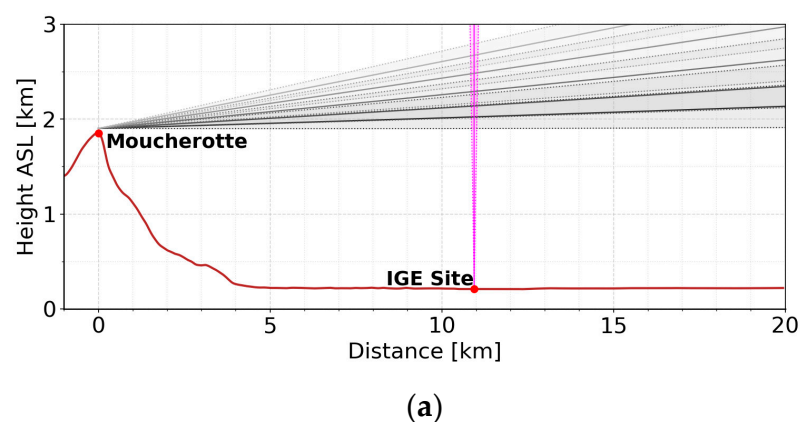


Figure 2. Cont.

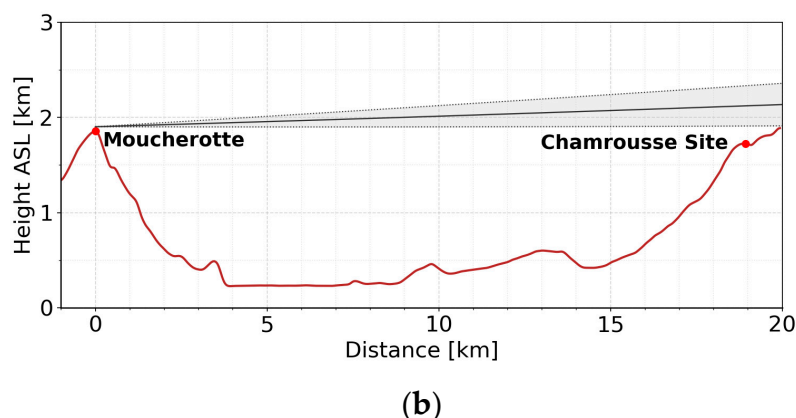


Figure 2. Vertical cross-sections illustrating terrain profile (red line) and radar beams (solid dark line indicates the center of the beam and grey shading indicates the 3 dB main lobe) between (a) the Moucherotte and the IGE sites, XRAD PPIs (0.6° , 1.2° , 2.0° , 3.0° , and 4.0°), and the MRR beam (solid pink line indicates the center of the beam and pink shading indicates the 3 dB main lobe); (b) the Moucherotte and Chamrousse sites, XRAD 0.6° PPI.

2.2. Datasets

The study spans from January 2020 to December 2023. Due to maintenance or other technical issues, concurrent observations for all instruments were not always available. Since the main objective of this study is to compare and validate different hydrometeor classifications between two instruments at a time, we opted to use all concurrent observations for each comparison considered. Consequently, the number of samples used varies among the different comparisons.

2.2.1. X-Band Weather Radar

Weather radar data were provided by the Météo-France operational X-band polarimetric Doppler radar located at Mt. Moucherotte, with an antenna altitude of 1913 m ASL [21]. XRAD observations consist of successive PPI scans recorded over a 5 min scanning cycle. Five of the six lowest elevation angles for this scanning cycle were utilized: 0.6° , 1.2° , 2.0° , 3.0° , and 4.0° (0.0° was excluded to avoid contamination from side lobes).

The XRAD operates with a range gate resolution of 240 m and a 3 dB beam width of 1.3° . The central beam height h at a distance r from the radar is given by [22]:

$$h = \sqrt{r^2 + (k_e R)^2 + 2rk_e R \sin(\theta)} - k_e R + H_0, \quad (1)$$

where R is the Earth's radius (6371 km), k_e is the Earth equivalent ratio factor for standard propagation conditions ($4/3$), θ is the antenna elevation angle, and H_0 is the antenna height (1913 m ASL). Given the short range between the radar and the area of interest (10 km), attenuation effects are considered negligible, which is consistent with previous studies showing that X-band radars typically experience significant path-integrated attenuation beyond 16–20 km under intense Mediterranean rainfall [23].

Radar data post-processing was performed using the Python ARM Radar Toolkit Py-ART [24], version 1.19.1, an open-source library designed for working with weather radar data. The variables used in the analysis include the equivalent radar reflectivity factor measured in the horizontal plane (hereafter Z_e) expressed in dBZ, the cross-correlation coefficient (ρ_{HV}) between the horizontally and vertically polarized returns, the differential phase (Φ_{dp}) in degrees, and the differential reflectivity (Z_{dr}) in dB [25–27]. From these variables, the specific differential phase (K_{dp}) is derived using the methodology proposed in [28]. Precipitation is only considered when Z_e exceeds 0 dBZ.

2.2.2. Micro Rain Radar

Data from a K-band (24.3 GHz) Doppler vertical profiler MRR, model MRR-2, manufactured by Metek, Meteorologische Messtechnik GmbH, Elmshorn, Germany, [29,30] were used. The MRR is installed at the IGE site and collocated with a disdrometer. The MRR was configured with a temporal resolution of 1 min and a vertical resolution of 100 m. It provides measurements across 30 vertical levels, ranging from 100 m to 3000 m above ground level (hereafter AGL). For analysis, the first two levels (100 and 200 m) are discarded to avoid ground contamination from side lobes. Thus, data from 300 m AGL to 3000 m AGL are used.

Post-processing of the MRR data is performed using the RaProM methodology [15,16], which enables hydrometeor classification across multiple altitudes. Precipitation is only considered when the radar reflectivity (Z_e) exceeds 0 dBZ for at least three consecutive minutes at a minimum of 10 distinct vertical levels, consistent with other similar studies [27–29]. Observations that do not meet this condition are excluded from the analysis.

The MRR radar constant was calibrated following the manufacturer’s methodology, as implemented in previous studies [31,32]. First, accumulated rainfall estimates over 30 min intervals were compared with concurrent measurements from the IGE automatic weather station rain gauge. Then, a calibration factor, M , was calculated as the median ratio between MRR-estimated and gauge-measured rainfall for these intervals. The obtained factor ($M = 0.18$) was then applied to correct the raw spectral reflectivity data from the MRR. After calibration, the RaProM processing was repeated on the adjusted data.

To facilitate comparison between the hydrometeor classification provided by the MRR and that obtained from the XRAD, an integration analysis was conducted to identify which MRR vertical levels intersected the XRAD beam volumes [33]. The contribution of each MRR level within the beam was weighted according to the beam’s vertical distribution, assuming a Gaussian-like decay in sensitivity from the middle of the radar beam [34]. Table 1 summarizes the analyzed MRR altitudes and the weights assigned to each level.

Table 1. Correspondence between XRAD PPI antenna elevations, height of the beam above sea level, MRR levels and their weight relative to the radar beam.

PPI Elevation [deg]	Beam Height [m ASL]	MRR Levels	Weight of Each Level
0.6	2032	16-17-18	0.29-0.39-0.32
1.2	2143	17-18-19	0.27-0.39-0.34
2.0	2292	19-20	0.47-0.53
3.0	2478	21-22	0.49-0.51
4.0	2665	22-23-24	0.25-0.38-0.37

Based on these weights, MRR reflectivity values are aggregated into a single weighted value representing the portion of the MRR column observed by the radar beam. In contrast, all hydrometeor classification values within the beam are considered equally (i.e., without weighting). All classification types are stored in an array to subsequently select the representative precipitation type.

To enable direct comparison with the radar, both the disdrometer and MRR data are temporally aggregated to 5 min intervals. For each interval, the median reflectivity is computed, and the corresponding hydrometeor classification types are stored in an array. Since different vertical levels are analyzed, up to a total of 15 classification values can be present within each 5 min interval. To determine a representative category, only the three most frequent hydrometeor types are considered for the final selection. Furthermore, the

second and third most frequent types are only retained if each account for more than 33% of the total occurrences within the 5 min period and across all analyzed levels.

2.2.3. Disdrometer

Three OTT Parsivel² disdrometers manufactured by OTT HydroMet, Kempten, Germany, [35,36] were utilized in this study, located at: the IGE research center (collocated with the MRR), Saint-Martin-d’Uriage (904 m ASL), and the Chamrousse AWS site (1730 m ASL). These instruments recorded measurements with a temporal resolution of 1 min, storing raw data for post-processing to derive the rainfall drop size distribution (DSD) and associated microphysical parameters. The DSD, which represents the number concentration of raindrops as a function of their diameter, is essential for the analysis of rainfall microphysics processes.

A quality control procedure was applied to remove non-meteorological artefacts in the DSD records, such as spurious detections caused by insects or instrumental noise, as described in previous studies [37–39]. To ensure data reliability, a minute of data was considered valid precipitation only if it belonged to a set of at least three consecutive minutes with more than 50 detected particles per minute. Minutes that did not meet this criterion were excluded from the analysis. The disdrometer data served two main purposes: (1) to validate the hydrometeor classification results at ground level at both the IGE and Chamrousse sites using the manufacturer’s software, and (2) to compare the measured reflectivity, Z , derived from the DSD [40–43] under liquid rainfall conditions. The reflectivity, Z , expressed in dBZ, is calculated using the following equation:

$$Z(\text{dBZ}) = 10\log_{10}\left(\sum_i N(D_i) \cdot D_i^6 \cdot \Delta D_i\right), \quad (2)$$

where D_i is the mean volume equivalent diameter of drop size class i (mm). The ΔD_i is the width of the diameter class (mm) and $N(D_i)$ is the drop concentration per unit volume and per unit diameter for each class i ($\text{m}^{-3} \text{mm}^{-1}$). The summation is performed for each drop size class i defined by the disdrometer. The drop concentration $N(D_i)$ for a given diameter i is calculated as [40]:

$$N(D_i) = \sum_{j=1}^n \frac{n_{ij}}{A_i \cdot V_i \cdot \Delta D_i \cdot \Delta t}, \quad (3)$$

where n_{ij} is the number of drops recorded in diameter class i and velocity class j . A_i is the effective sampling area of the disdrometer for each diameter class i , Δt is the sampling interval time (60 s), ΔD_i is the width of the diameter class. Finally, unlike [40], we do not use the disdrometer fall velocities measure. Instead, we use V_i , the theoretical fall velocity for each diameter i according to the following terminal velocity expression [44,45], where V is given in m/s and D is expressed in mm:

$$V = 9.5 \left(1 - e^{-0.6D}\right). \quad (4)$$

To enable temporal comparison with the XRAD, the disdrometer data were aggregated over 5 min intervals. For each interval, the median reflectivity was computed, the total precipitation accumulation was summed, and all values related to the precipitation type were stored in an array. This procedure allowed a subsequent comparison with the XRAD and MRR, selecting the best match according to the usual principle of fuzzy verification methods [46].

The disdrometer at IGE was operational during the same periods as the MRR, with the data analyzed corresponding to the period from 2020 to 2023. The Chamrousse and Saint-Martin-d’Uriage disdrometers have been operating continuously since January 2021.

2.2.4. Automatic Weather Station and Temperature Profile

Data from 11 surface AWSs are used in this study, including 10 stations operated by Météo-France and one managed by IGE. The Météo-France stations are distributed across the Grenoble valley, providing comprehensive spatial coverage across different elevations (from the valley floor at 220 m ASL up to 2240 m ASL) and located within a maximum distance of about 25 km from the IGE site. These AWS datasets include hourly averages of air temperature and relative humidity (measured at 1.5 m AGL), wind speed and direction (measured at 10 m AGL), and accumulated precipitation. Table 2 lists the Météo-France AWS used, including their name, geographical coordinates, altitude, and distance from the IGE site. The IGE station, located at the IGE research center on the valley floor, records the same meteorological parameters but at a higher temporal resolution of 10 min.

Table 2. Name, latitude, longitude, altitude, and distance from the IGE site of Météo-France AWSs used.

AWS Name	Latitude [deg]	Longitude [deg]	Altitude [km ASL]	Distance from IGE [km]
Grenoble-CEA	45.21	5.68	0.22	6.5
Grenoble-LVD	45.22	5.85	0.22	7.3
Theyss	45.30	6.00	0.62	21.9
Saint-Hilaire SAPC	45.31	5.88	0.99	16.5
Col de Porte Nivose	45.29	5.77	1.32	11.2
Pipay SAPC	45.27	6.02	1.58	21.4
Le Gua Nivose	45.01	5.59	1.65	24.1
Chamrousse	45.13	5.88	1.73	11.7
Saint-Hilaire Nivose	45.31	5.86	1.76	15.5
Aigle-ton Nivose	45.23	6.04	2.24	22.0

2.2.5. AROME Forecast

Data from the operational high-resolution numerical weather prediction model AROME [47,48] was utilized in this study. AROME provides hourly forecasts with a horizontal resolution of approximately 1.3 km. Temperature data were extracted at the following predefined heights AGL: 2, 10, 20, 35, 50, 75, 100, 150, 200, 250, 375, 500, 625, 750, 875, 1000, 1125, 1250, 1375, 1500, 1750, 2000, 2250, 2500, 2750, and 3000 m. The reference surface for these heights corresponds to the AROME model orography, which is approximately 257 m ASL for the Grenoble area.

2.3. Methodology

Our methodology for validating the classification methods proceeds in three distinct steps: first, the MRR classification is checked against disdrometer observations, which is considered the original ground-truth reference, following the approach used in [15,49]; however, note that no modifications are performed to the MRR classification during this step. Then, the MRR classification is considered the reference in subsequent comparisons with XRAD-based classification methods. This approach allows for a more consistent comparison than would be possible in an XRAD to disdrometer data comparison given the fact that both MRR and XRAD share common observation volumes, whereas the disdrometer at Chamrousse site is located approximately 397 m below the radar center beam level, introducing a spatial mismatch. Finally, XRAD is compared directly with the disdrometer.

This final step serves to evaluate and contrast its performance against the original ground reference and an independent dataset used in the initial methodology comparison.

2.3.1. Hydrometeor Classification Methods

Different methods were employed for hydrometeor classification depending on the instrument. Disdrometer data were classified using the manufacturer's proprietary software, here referred to simply as the disdrometer classification. For the MRR, the RaProM program was applied as previously explained (hereafter, MRR classification).

The XRAD hydrometeor classification was conducted using three main schemes, based on the polarimetric variables and with temperature as an optional parameter, which finally yielded a total of 10 different variants. The first two schemes are part of the Colorado State University (CSU) Radar Tool Kit [50] and the third is based on random forest machine learning.

The first CSU scheme, called *csu_fhc_summer* [10,51] (hereafter HCS), was specifically designed for warm weather, but can also be applied to cooler conditions. It uses radar variables (Z_e , Z_{dr} , ρ_{HV} , and K_{dp}) for each individual radar pixel as input, with temperature as an optional parameter. HCS classifies hydrometeors into ten categories. This methodology operates pixel-by-pixel in the radar data, independent of the surrounding pixels.

The second CSU scheme, called *run_winter* (hereafter HCW), is described in [11] and was designed for cool weather conditions where the freezing level is often contained in the radar volume. It uses as input data polarimetric radar variables and, optionally, a temperature profile. First, the method attempts to determine the freezing level, and then classifies hydrometeors into seven categories. If temperature is available, it is used to assess the possibility of frozen precipitation below the freezing level. Unlike HCS, HCW requires all PPI data from the radar to evaluate hydrometeor classification, so it cannot be applied to isolated individual radar pixels.

This study also introduces a new hydrometeor classification methodology based on machine learning. A random forest classifier [52] was selected because it is a decision-tree-based approach that allows an analysis of the criteria used for classification. Thus, results can be interpreted more easily in physical terms. The input features were limited to a maximum of five (four polarimetric radar variables and temperature), making random forest an appropriate choice for this context. The random forest algorithm works by constructing an ensemble of decision trees, each trained on a random subset of the data and features, and then combining their predictions through majority voting.

The random forest classifier was implemented using Scikit-learn [53], version 1.6.1. Hyperparameter tuning was performed using a grid search combined with five-fold cross-validation on the training set. The objective metric was overall accuracy, with secondary checks on class-balanced performance. The tested ranges included tree depth (3–10), minimum samples per split (5–30), and impurity decrease thresholds (0.0001–0.01). The configuration reported above yielded the best validation accuracy. The total number of decision trees was set to 100, the splitting criterion was Gini impurity, and the maximum tree depth was limited to five levels. The minimum number of samples required to split an internal node was fixed at 15, and the minimum number of samples required at a leaf node was also set to 15. A node was split only if the split resulted in an impurity decrease greater than or equal to 0.001. Finally, the class weights were adjusted to the 'balanced' mode to account for class imbalance. For training and testing, the dataset was randomly divided into a training set and a testing set, reserving 20% of the data for testing. Additionally, to address concerns regarding temporal autocorrelation within consecutive minutes of precipitation, a blocked validation approach was implemented. Five days were randomly selected and excluded from the training set to form the test set, thereby

reducing the influence of event-specific temporal structures. To make the results more robust and independent of the specific days selected, this procedure was repeated 100 times, each time selecting five different days as the test set. Details of the predictor set, target variable, pre-processing steps, QC thresholds, and RF hyperparameters are summarized in Table A1 (Appendix A).

The algorithm classifies hydrometeors into three categories, using the MRR classification as the target. A total of five different schemes based on this random forest approach were evaluated: three using different temperature profiles; one using only the vertical temperature profile, with Z_e as a mask for precipitation presence; and one using all radar variables without temperature information. These last two schemes were included to assess the added value of combining polarimetric radar data with temperature, as opposed to using these data sources separately.

Since each hydrometeor classification method (Disdrometer, MRR, HCS, HCW, RF) has its own unique set and number of categories, a unified classification was adopted for intercomparison. Table 3 shows the mapping between the original classification and the three final categories considered. For simplicity, these categories correspond to the predominant hydrometeor phase, i.e., solid, liquid, and mixed, hereafter referred to as rain, mixed, and snow. Although a simplified classification based on predominant phase (rain, mixed, snow) was adopted, the mixed category is especially sensitive to uncertainty near the melting layer. The absence of a universal threshold for defining mixed precipitation introduces ambiguity that can affect the coherence of intercomparisons. Nevertheless, we consider that these classifications are comparable enough to allow the intercomparison study to be carried out.

Table 3. Hydrometeor categories for each classification method (HCS, HCW, Disdrometer, and MRR) and the simplified classification considered in this study (rain, mixed, and snow).

HCS	HCW	Disdrometer	MRR	This Study
Drizzle		Drizzle	Drizzle	
Rain Big drops	Rain	Drizzle with rain Rain	Rain	Rain
Wet/melting snow	Wet snow	Rain drizzle with snow	Mixed	Mixed
Ice crystals Aggregates Vertically aligned ice Low-density graupel	Ice crystals Aggregates Plates Dendrites	Snow Snow grains	Snow	Snow
High-density graupel Hail	Frozen precipitation	Soft hail Hail	Hail	Not considered

The Hail category was excluded from the analysis due to its negligible occurrence during the observation period and the lack of consistent detection across all instruments.

Only those time intervals where both instruments being compared detected some type of precipitation were considered. Minutes classified as ‘no-detection’ or ‘unknown’ were excluded from the analysis. As previously mentioned, the disdrometer and the MRR data were temporally aggregated to 5 min intervals, with the MRR analysis also considering specific vertical levels. All identified hydrometeor types within each 5 min period were stored in a vector for comparison. For the MRR, the three most frequent categories within that interval are considered, if each reaches a minimum weight of 33% of the total occurrences. The comparison is then performed by checking whether any of the categories meeting this condition coincide between the two instruments. If no match

is found, the category with the highest overall weight is selected for the comparison. The 33% threshold was selected because it corresponds to the chance-level share for three classes (rain, mixed, snow), ensuring that only categories with substantial representation are included. This prevents spurious matches from low-frequency labels. Empirically, ~75% of intervals retained the modal category after aggregation.

2.3.2. Temperature Profiles

To characterize the vertical temperature profile in the region of study, three different approaches are utilized. The first one, based on the AROME model forecast column over the IGE site, builds a temperature profile that interpolates the temperature model levels linearly in space and time to match the temporal intervals and altitudes required for comparison with XRAD observations.

The second approach consists of using temperature in-situ measurements from AWS stations at different altitudes, following an approach similar to [54], assuming near-surface temperature homogeneity across the valley floor. This hypothesis was assessed using multiple low-altitude stations, as in [55]. Specifically, the median ground-level temperature was computed from three stations in Grenoble city (IGE, Grenoble-CEA, and Grenoble-LVD). Four additional stations located between 1578 m and 1756 m ASL (Chamrousse, Saint-Hilaire-Nivose, Le Gua-Nivose, and Pipay Sapc) were combined to define a single level at approximately 1722 m ASL. This approach yields a simplified vertical profile spanning from 220 m to 2240 m ASL, using six altitude levels: 220 m (median), 615 m, 990 m, 1325 m, ~1720 m (median), and 2240 m. Temperatures among these levels were linearly interpolated, and above 2240 m, extrapolated using a standard atmosphere environmental lapse rate of 6.5 °C/km. Finally, hourly temperature data were interpolated to a 5 min resolution to match radar observations.

The third method is based on finding the freezing level height derived from the MRR vertical hydrometeor classification, assumed to occur near the 0°C level. Using 10 min temperature measurements from the IGE AWS, collocated with the MRR, the vertical temperature gradient is estimated at each time step, allowing the reconstruction of a vertical temperature profile. The profile is derived as follows: for every five averaged MRR profiles, the freezing level height is extracted. However, freezing level estimates often display high variability due to signal noise. To mitigate this, several filters are applied. First, daily freezing level values exceeding one standard deviation from the global dataset are excluded to remove potential outliers. Second, when no freezing level is detected, a median temperature lapse rate is assigned, either the daily median or, if unavailable, the annual median temperature lapse rate of −6.8 °C/km, calculated from the MRR data. Finally, if the resulting temperature lapse rate falls outside the interquartile range, it is replaced with a value within the range, as such extremes are considered erroneous. Table A2 (Appendix B) summarizes all classification methods.

2.3.3. Verification Metrics

To characterize and compare the performance of the different hydrometeor classification methods, a contingency table (confusion matrix) is used to compute the associated verification scores [56–58], described below.

Accuracy is defined as the proportion of total correct predictions relative to the total number of samples. It ranges from 0 (no correct predictions) to 1 (perfect classification):

$$Accuracy = \frac{\sum_{i=1}^N c_{ii}}{T}, \quad (5)$$

where c_{ii} is the number of corrected predictions for class i , N is the total number of classes and T the total number of samples.

Heidke skill score (HSS) [59] evaluates the skill of the classification compared to random chance:

$$HSS = \frac{\sum_{i=1}^N c_{ii}T - \sum_{i=1}^N C_i R_i}{T^2 - \sum_{i=1}^N C_i R_i}, \quad (6)$$

where R_i is the total number of observed instances of class i and C_i the total number of predicted instances as class i . HSS ranges from 0 (random classification) to 1 (perfect classification).

Gerrity skill score (GSS) [60], evaluates the performance of precipitation classification across the three categories considered (rain, mixed, and snow), assigning different weights to misclassifications so that errors between adjacent classes (e.g., snow classified as mixed) are penalized less severely than errors between distant classes (e.g., snow classified as rain). GSS ranges from 0 (no skill, equivalent to random classification) to 1 (perfect classification). GSS is computed according to:

$$GSS = \frac{\sum_{i=1}^N \sum_{j=1}^N c_{ij} S_{ij}}{T}, \quad (7)$$

where S_{ij} is a parameter that assigns a weight based on the difference between classes, so that misclassifications between adjacent classes are penalized less severely than those between non-adjacent classes.

Precision reports the proportion of correctly predicted instances among all instances predicted as class i :

$$Precision_i = \frac{c_{ii}}{C_i}, \quad (8)$$

Recall is the proportion of correctly predicted instances among all actual instances of class i :

$$Recall_i = \frac{c_{ii}}{R_i}, \quad (9)$$

(note that precision is also known as success ratio and recall as probability of detection according to [61]).

F1-score provides a single score that balances the recall and the precision:

$$F1_i = 2 \frac{Precision_i \cdot Recall_i}{Precision_i + Recall_i}, \quad (10)$$

and finally, Frequency Bias indicates whether a class tends to be overpredicted (value greater than 1) or underpredicted (value lower than 1):

$$Frequency\ Bias_i = \frac{C_i}{R_i}. \quad (11)$$

3. Results

3.1. Overview of Selected Events

An example of the events examined is presented in Figure 3, corresponding to 1 December 2023. At the beginning of the event, precipitation occurred as rain above 2.0 km ASL and after 19:00 UTC, and snow reached the surface at 0.2 km ASL at the IGE site, illustrating a transition phase observed by all instruments.

As illustrated in Figure 3a, this case corresponds to a stratiform precipitation event characterized by a distinct bright band around 2.52 km ASL at the beginning of the day. This bright band gradually descended, reaching the surface at the IGE site between approxi-

mately 13:00 and 19:00 UTC. Additionally, it is visually apparent that all three instruments—XRAD, MRR at 2.0 km ASL, and the disdrometer at 1.7 km ASL—detect similar reflectivity values and evolution, indicating proper calibration of the MRR and good agreement among the instruments.

Similarly, at lower altitudes (MRR at 0.5 km ASL and the disdrometer at 0.2 km ASL), a consistent relationship between the reflectivity is observed, until the bright band reaches the surface. At that point, the disdrometer records significantly higher reflectivity than the MRR. This difference can be primarily explained by the 300 m vertical difference between the lowest MRR measurement level and the disdrometer location.

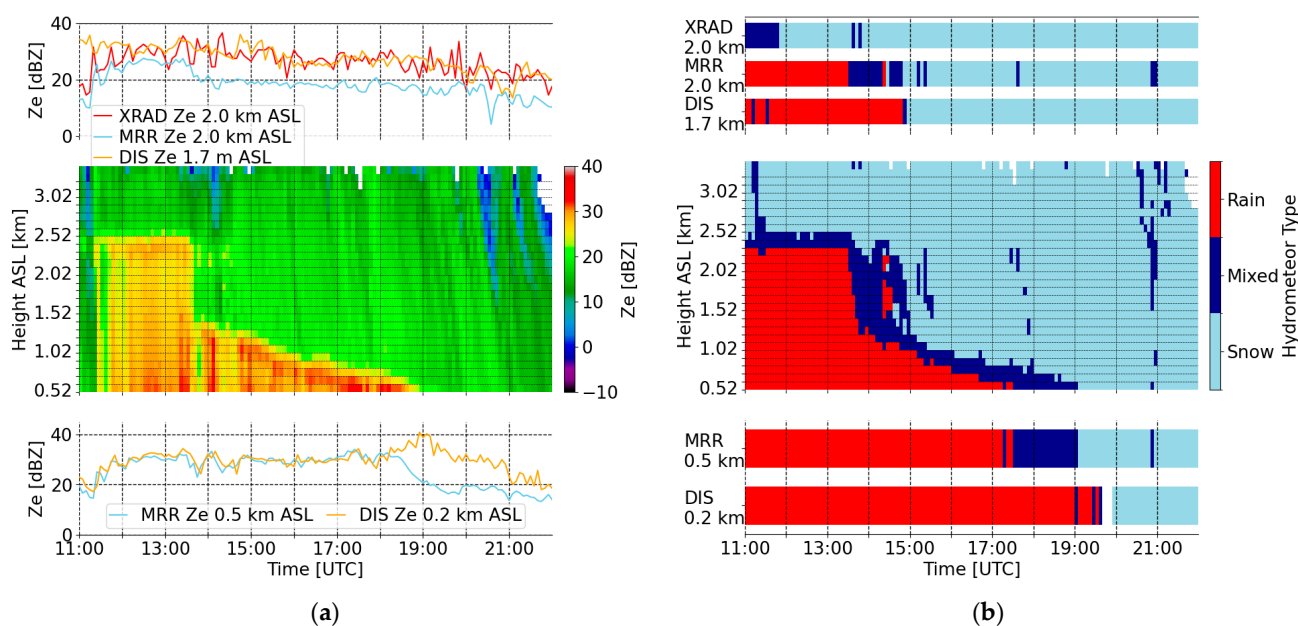


Figure 3. 1 December 2023 event, showing a transition from liquid to solid precipitation from 2500 m ASL down to ground level (220 m ASL). (a) Top panel: Time series of equivalent radar reflectivity factor Z_e [dBZ] from the XRAD (red) and MRR (blue) at 2020 m ASL, and from the Chamrousse disdrometer (orange) at 1720 m ASL. Middle: Vertical profile of reflectivity from the MRR from 520 m to 3220 m ASL. Bottom panel: Time series of Z_e [dBZ] from the MRR at 520 m ASL (blue) and the IGE disdrometer at 220 m ASL (orange). (b) Same panels as (a), but for hydrometeor classification (Rain, Mixed, Snow) using the HCS_AWS for the XRAD.

Figure 3b shows the hydrometeor classification, and, according to the MRR, the phase change aligns well with the bright band seen in Figure 3a. Hydrometeor types classified by the MRR and the disdrometer agree well, detecting the phase transition almost simultaneously both in height and at the surface. In contrast, the XRAD classification shows less agreement with the MRR at higher altitudes, as it records more mixed or snow precipitation during the first hours, while the MRR still detects rain. Despite this discrepancy, once the MRR detects snow, the XRAD also identifies snow.

3.2. Disdrometer vs. MRR

Figure 4 presents the contingency table comparing hydrometeor classifications based on disdrometers and the MRR, using 44,387 min of concurrent observations. This includes data from three measurement sites: the IGE site (with 19,142 min), the Saint-Martin-d’Uriage site (10,675 min), and the Chamrousse site (14,570 min).

As shown in Figure 4, there is a strong agreement (demonstrated by the high values on the diagonal or the contingency table), indicating a good correspondence between the classifications from both instruments. Table 4 shows the statistical values of Figure 4, with

an Accuracy of 86%, HSS of 0.68, and a GSS of 0.82, suggesting substantial agreement between two instruments. The performance metrics by class reveal high Precision and Recall for most hydrometeor types. However, the mixed category shows poorer results, particularly in terms of Precision, with a value of 0.08 and a Frequency Bias of 10.45. This indicates that the MRR overestimates the mixed category, while the disdrometer tends to classify these cases as rain. This discrepancy may arise from the fact that, at the IGE site, the MRR data used in the comparison corresponds to its third vertical level, located approximately 300 m above the surface disdrometer. This vertical separation may introduce systematic classification differences due to possible phase changes occurring within that layer. At the Chamrousse and Saint-Martin-d'Uriage sites, the spatial separation between the MRR (located at IGE) and the disdrometer (located ~10 km away) must also be considered, as local meteorological variability can affect the observed hydrometeor types. Furthermore, differences may also stem from the intrinsic nature of the instruments: the disdrometer provides point measurements at the surface, whereas the MRR retrieves data over a vertical volume. This difference in spatial and temporal sampling can lead to the disdrometer underestimating mixed-phase precipitation while the MRR overestimates it. Therefore, the high frequency bias observed for this category should not necessarily be interpreted as an error, but rather as a consequence of the inherent characteristics of the instruments.

Despite these spatial and vertical differences, the matrix indicates good agreement between both instruments for rain and snow, while discrepancies in mixed-phase data remain. Nevertheless, this level of consistency can be considered satisfactory for the purposes of intercomparison and supports the overall reliability of their respective classification algorithms.

Table 4. Statistical values of Accuracy, HSS, Precision, Recall, F1-score, and Frequency Bias for the hydrometeor classification results comparing disdrometer vs. MRR.

Disdrometer vs. MRR	
Accuracy [%]	86
HSS	0.68
GSS	0.82
Precision	
Rain	0.99
Mixed	0.08
Snow	0.87
Recall	
Rain	0.86
Mixed	0.79
Snow	0.86
F1-score	
Rain	0.92
Mixed	0.14
Snow	0.86
Frequency Bias	
Rain	0.87
Mixed	10.45
Snow	0.99

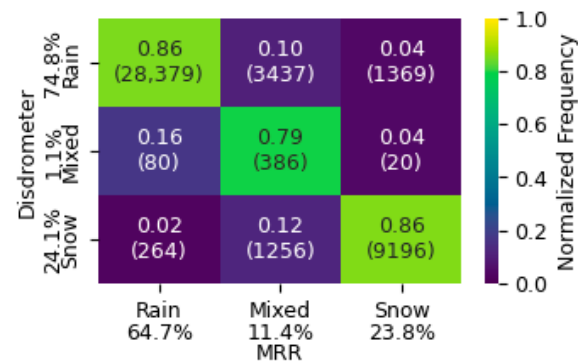


Figure 4. Contingency table comparing hydrometeor classifications based on disdrometer and MRR data. Values in each cell indicate the proportion relative to the disdrometer category (color-coded) and the absolute count (in parentheses). Percentages for each hydrometeor category are also indicated: for the MRR, along the *x*-axis, and for the disdrometer, along the *y*-axis, representing the relative frequency of each category for each instrument.

3.3. MRR vs. XRAD

In the following section, the results of the different sub-methods for hydrometeor type classification based on the HCS and HCW are analyzed with respect to the classification obtained from the MRR. The analysis includes all available PPI elevation angles, with a total of 34,685 min processed, summing all minutes across the selected levels. Figure 5 presents the contingency tables for the two best-performing methods: HCS_AWS and HCS_AROME, while Table 5 summarizes the statistical results of these methods, also including HCW_NTemp, HCS_NTemp, and HCS_FL.

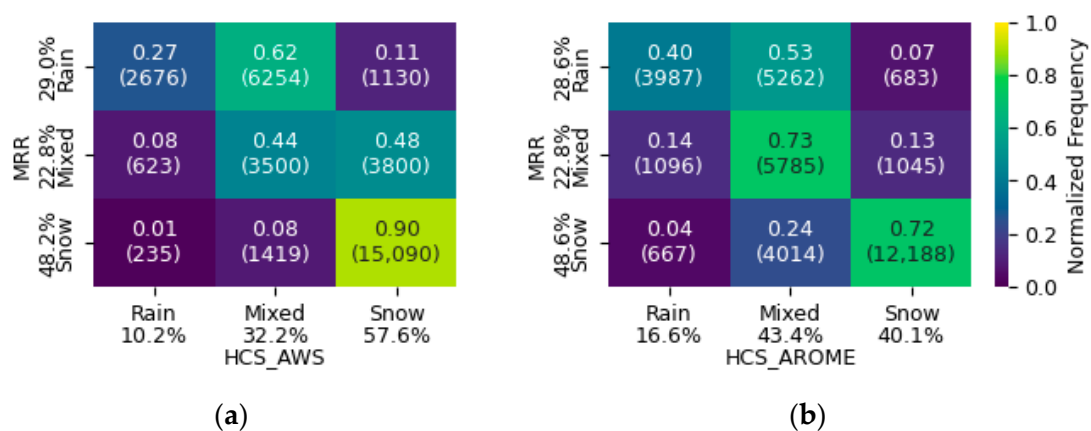


Figure 5. Contingency table comparing hydrometeor classifications based on MRR and (a) HCS_AWS, (b) HCS_AROME data. Values in each cell indicate the proportion relative to the MRR category (color-coded) and the absolute count (in parentheses). Percentages for each hydrometeor category are also indicated: for the (a) HCS_AWS and (b) HCS_AROME, along the *x*-axis, and for the MRR, along the *y*-axis, representing the relative frequency of each category for each instrument.

As shown in Table 5, the best classification performance is achieved with HCS_AROME, reaching an Accuracy of 63%, followed by HCS_AWS at 61%, and HCS_FL at 55%. Similarly, the highest GSS and HSS scores are also obtained by HCS_AROME. In contrast, methods that do not use temperature information (HCS_NTemp and HCW_NTemp) show limited agreement with the reference classification. These approaches exhibit a strong tendency to classify hydrometeors as mixed, resulting in a high Frequency Bias for this class and very low frequencies for rain and snow. This behavior negatively affects class precision and skill scores, with HSS and GSS values of 0.09 and 0.19 for HCS_NTemp, and 0.10 and 0.19

for HCW_NTemp, respectively. Such low values indicate that their performance is only marginally better than random classification. Consequently, overall Accuracy is poor, 35% for HCS_NTemp and 44% for HCW_NTemp. These results clearly demonstrate that incorporating temperature information substantially improves classification accuracy, highlighting the importance and benefits of using a temperature profile in hydrometeor classification.

Table 5. Same as Table 4, but showing classifications comparing MRR vs. HCS_NTemp, MRR vs. HCW_NTemp, MRR vs. HCS_AWS, MRR vs. HCS_FL, and MRR vs. HCS_AROME.

	HCS_NTemp	HCW_NTemp	HCS_AWS	HCS_FL	HCS_AROME
Accuracy [%]	35	44	61	55	63
HSS	0.09	0.19	0.37	0.33	0.44
GSS	0.10	0.19	0.43	0.42	0.48
Precision					
Rain	0.45	0.75	0.76	0.66	0.69
Mixed	0.28	0.31	0.31	0.27	0.38
Snow	0.60	0.68	0.75	0.84	0.88
Recall					
Rain	0.24	0.16	0.27	0.43	0.40
Mixed	0.81	0.83	0.44	0.59	0.73
Snow	0.17	0.41	0.90	0.60	0.72
F1-score					
Rain	0.31	0.27	0.39	0.52	0.51
Mixed	0.42	0.45	0.37	0.37	0.50
Snow	0.27	0.51	0.82	0.71	0.79
Frequency					
Bias					
Rain	0.53	0.22	0.35	0.66	0.58
Mixed	2.85	2.69	1.41	2.20	1.9
Snow	0.29	0.60	1.20	0.72	0.82

3.4. Machine Learning Classification

The next section presents the results of the five random forest hydrometeor classification methods. A training dataset consisting of 34,685 samples and a test dataset of 6961 samples were constructed. Figure 6 shows the results obtained from the model trained on the AROME profile (RF_AROME), which was also tested using temperature profiles from AWS.

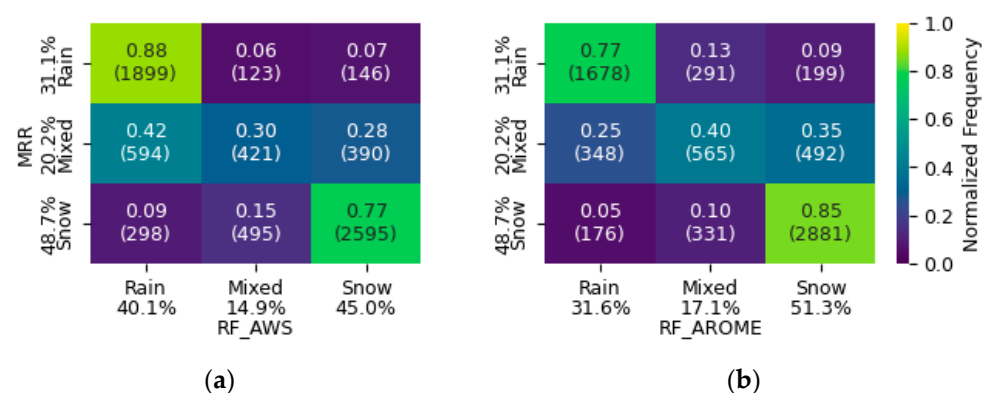


Figure 6. Contingency table comparing hydrometeor classifications based on MRR and (a) RF_AWS, (b) RF_AROME data. Values in each cell indicate the proportion relative to the MRR category (color-coded) and the absolute count (in parentheses). Percentages for each hydrometeor category are also indicated: for the (a) RF_AWS and (b) RF_AROME, along the x-axis, and for the MRR, along the y-axis, representing the relative frequency of each category for each instrument.

Table 6 summarizes the statistical results corresponding to Figure 6 and the rest of the random forest classification methods, RF_FL, RF_Temp and RF_NTemp. We can observe that all statistical values have improved with this new classification approach. The best performance is obtained when using the RF_AROME, with an Accuracy of 74% and an HSS of 0.57, indicating a substantial level of agreement. When the model was trained using the blocked validation approach, the median values across 100 iterations were very similar to those obtained with the original random 80/20 split. This confirms that the classification performance is robust and not significantly affected by temporal autocorrelation in the data. The results obtained with temperature profiles from AWS and the MRR also show improved performance, with Accuracy of 71% and 70% and HSS values of 0.53 and 0.50, respectively. Regarding the other statistical measures, the rain and snow classes are generally well classified, showing robust scores, while the mixed class remains the most challenging to identify correctly and is underestimated in this case, with a frequency bias of 0.84 and low precision and recall values of 0.48 and 0.40, respectively, for the RF_AROME model.

Table 6. Same as Table 5, but showing classifications comparing MRR vs. RF_Temp, MRR vs. RF_NTemp, MRR vs. RF_AWS, MRR vs. RF_FL, and MRR vs. RF_AROME.

	RF_Temp	RF_NTemp	RF_AWS	RF_FL	RF_AROME
Accuracy [%]	71	51	71	70	74
HSS	0.54	0.20	0.53	0.50	0.57
GSS	0.63	0.22	0.65	0.60	0.66
Precision					
Rain	0.77	0.44	0.68	0.74	0.76
Mixed	0.41	0.32	0.41	0.36	0.48
Snow	0.82	0.60	0.83	0.77	0.81
Recall					
Rain	0.72	0.49	0.88	0.68	0.77
Mixed	0.49	0.23	0.30	0.26	0.40
Snow	0.79	0.63	0.77	0.90	0.85
F1-score					
Rain	0.75	0.46	0.77	0.71	0.77
Mixed	0.45	0.27	0.34	0.30	0.44
Snow	0.80	0.62	0.80	0.83	0.83
Frequency					
Bias					
Rain	0.94	1.09	1.29	0.91	1.02
Mixed	1.19	0.72	0.74	0.73	0.84
Snow	0.96	1.06	0.92	1.17	1.05

Finally, we can observe that using only the vertical temperature profile, RF_Temp, for classification also leads to good results, with a global Accuracy of 71% and a GSS of 0.63. These results clearly demonstrate the dominant role of temperature in hydrometeor classification. While polarimetric radar variables showed a relatively small importance in distinguishing precipitation types, their inclusion resulted in an increased overall Accuracy of 74%. On the other hand, it should be noted that using only radar data without the vertical temperature profile yields an Accuracy of 51% and an HSS of 0.20, providing clear evidence for the need to include the temperature profile. Nevertheless, these results are still better than the HCW_NTemp classification, which did not use either temperature for classification. Finally, it can be stated that, although the temperature profiles obtained from

RF_AWS and RF_FL methods yield slightly lower classification performance, they can still be considered sufficiently accurate and applicable for practical use.

To better understand the physical processes governing the random forest model's decisions, we analyzed the feature importances and the median classification criteria of the first and second levels across the 100 constituent trees. As anticipated for precipitation phase classification, temperature is the most influential variable, exhibiting a dominant importance score of 0.87 (on a 1-point scale). It is followed by Z_e and Z_{dr} , with importance scores of 0.05 and 0.04, respectively. Both variables are strongly associated with the precipitation phase: Z_e tends to be lower for snow and higher in the mixed phase associated with the bright band, while Z_{dr} indicates the degree of particle flattening, a typical feature of liquid drops. Finally, K_{dp} and ρ_{HV} , with lower scores of 0.02 and 0.01, play a less significant role in the overall model structure. To specifically probe the mixed-phase regime, we conducted a targeted experiment restricting the evaluation to temperatures between $-2\text{ }^{\circ}\text{C}$ and $+6\text{ }^{\circ}\text{C}$, where snow, mixed, and rain can coexist. As expected for this transition zone, the overall accuracy decreases to $\sim 65\%$ due to the increased ambiguity near phase boundaries. Within this range, temperature remains the most influential predictor, but with a reduced importance score of 0, reflecting the greater role of non-thermal signatures in mixed-phase conditions. Z_e is the second most informative variable, with an importance of 0.12, which is consistent with enhanced bright-band effects. The remaining polarimetric variables retain importance values comparable to the full-sample analysis.

The analysis of the median classification criteria across the first level of each decision tree of the random forest confirms the primary role of temperature and reveals the physical nature of the resulting splits. In the first level, 41% of the trees utilize temperature for the initial split, using a median threshold of $-0.5\text{ }^{\circ}\text{C}$. This fundamental division classifies temperatures below this threshold as snow and higher values as rain, effectively separating the solid and liquid precipitation regimes. The second classification criteria involve Z_e , which is used in 24% of the trees, dividing the dataset at approximately 29 dBZ by classifying lower values as snow and higher values as mixed. This high reflectivity threshold is consistent with the characteristics of the bright band associated with the mixed phase. Similarly, Z_{dr} splits the dataset at 24% of the trees at a median threshold of 0.9 dBZ, separating values below this into the snow category and higher values into rain, reflecting the expected higher degree of particle flattening in rain. K_{dp} is utilized in 11% of the trees and divides the dataset into rain values below 0.23 and mixed values above. The criteria observed in the second level are highly similar to those in the first, with temperature remaining the most common variable for splitting 42%, reinforcing the stability and robustness of these physical classification rules throughout the model. Furthermore, the robustness of the classification method is highlighted by the performance of the individual estimators: the median Accuracy of each individual tree is 70%, and the standard deviation across the 100 trees is only 0.04 points.

3.5. Disdrometer vs. XRAD

Finally, RF_AROME and HCS_AROME are evaluated against the disdrometer instrument, located at a higher elevation in Chamrousse. It should be noted that the radar beam and the disdrometer are separated by a vertical distance of approximately 397 m, which may influence the observed hydrometeor phase due to temperature and microphysical gradients. Figure 7 shows the contingency table of the classification using HCS_AROME and RF_AROME. In this case, since the XRAD data correspond to the Chambrousse site and were not included in the training dataset from the IGE site, the evaluation represents an independent test. The RF_AROME model was trained using all data from the IGE site.

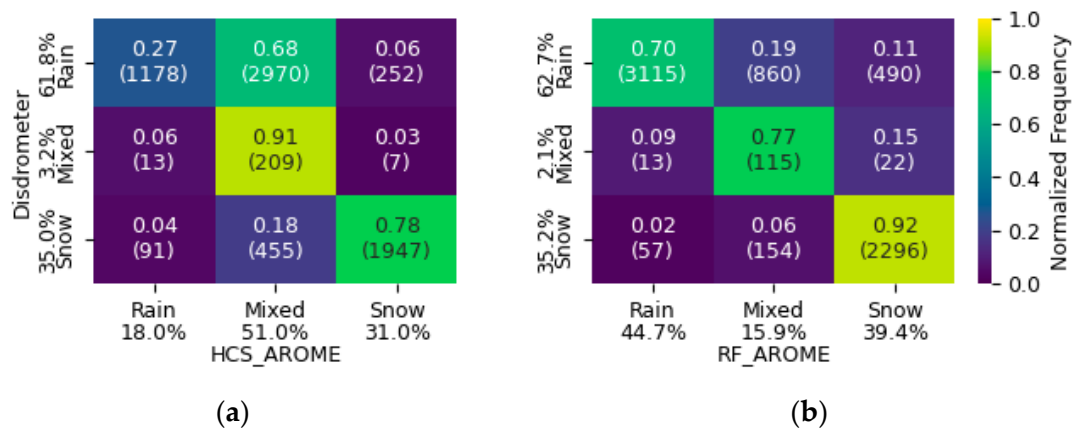


Figure 7. Contingency table comparing hydrometeor classifications based on disdrometer and (a) HCS_AWS, (b) RF_AROME data. Values in each cell indicate the proportion relative to the disdrometer category (color-coded) and the absolute count (in parentheses). Percentages for each hydrometeor category are also indicated: for the (a) HCS_AWS and (b) RF_AROME, along the x-axis, and for the disdrometer, along the y-axis, representing the relative frequency of each category for each instrument.

Table 7 shows the verification scores corresponding to Figure 7, based on 7122 min. The scores indicate a significant improvement in the classification provided by RF_AROME compared to HCS_AROME. The overall Accuracy increases from 47% to 78%, and all other scores also show a clear improvement. This is particularly important for the RF_AROME method, as the other methods tended to classify a large proportion of hydrometeors as Mixed, resulting in a Frequency Bias of 15.87 for the HCS_AROME model, which was reduced to 7.53 in the RF_AROME model.

Table 7. Same as Table 6, but showing disdrometer vs. HCS_AROME and disdrometer vs. RF_AROME.

	HCS_AROME	RF_AROME
Accuracy [%]	47	78
HSS	0.30	0.61
GSS	0.48	0.74
Precision		
Rain	0.92	0.98
Mixed	0.06	0.10
Snow	0.88	0.82
Recall		
Rain	0.27	0.70
Mixed	0.91	0.77
Snow	0.78	0.92
F1-score		
Rain	0.41	0.81
Mixed	0.11	0.18
Snow	0.78	0.86
Frequency Bias		
Rain	0.29	0.71
Mixed	15.87	7.53
Snow	0.88	1.12

It is also important to mention the difference in HSS and GSS scores for each method: HCS_AROME achieves values of 0.30 and 0.48, while the RF_AROME model reaches 0.61 and 0.74, respectively, indicating a substantial improvement. The high GSS value is particularly significant, as it indicates that most classification errors occur between adjacent classes. Consistent with this, we performed a temperature-based sensitivity analysis relative to the melting layer: when excluding the temperature transition range (-2°C to $+6^{\circ}\text{C}$), overall accuracy increases to 92%, with rain/snow frequency bias close to 1. However, within that range, accuracy drops to 68% and snow precision to 0.38, underscoring the challenge of mixed-phase conditions. This behavior can be partly explained by the 397 m vertical separation between the radar beam and the disdrometer.

4. Discussion

The results of this study confirm the feasibility of improving hydrometeor phase classification using remote sensing observations combined with machine learning techniques, specifically by integrating observations from MRR and an operational X-band polarimetric radar. Validation against independent instruments, such as disdrometers, demonstrates the robustness of the MRR-based classification RaProM [15,16], which was subsequently used as a reliable reference for assessing the performance of the operational XRAD schemes. This approach aligns with previous studies that ensure consistency between ground-based and radar-based observations. For instance, [62] validated radar-based hydrometeor classification using surface observations, while [33] emphasized the importance of cross-validation with disdrometer data to improve classification reliability.

A key finding of this study is that the two radar systems demonstrate contrasting requirements for hydrometeor classification. The RaProM methodology enables the MRR to classify hydrometeors without relying on external temperature profiles, achieving an Accuracy of 86% when compared to disdrometer observations. In contrast, XRAD shows a strong dependency on accurate vertical temperature information. Without temperature input, RF_NTemp classification Accuracy drops significantly to 51%, whereas incorporating temperature data, RF_AROME improves performance to 74% when benchmarked against MRR classification. This dependency aligns with previous research emphasizing the importance of thermodynamic profiles in operational radar-based hydrometeor classification [11,63]. The likely explanation for this difference lies in the radar capabilities: operational radars such as XRAD do not directly measure fall velocity, a key parameter for distinguishing hydrometeor phases, while the MRR does.

It is also noteworthy that a classification approach using only reflectivity and temperature, without polarimetric variables, achieves 71% Accuracy relative to the MRR. This classification performs three percentage points lower than the model incorporating both polarimetric data and temperature. While this represents a modest gain in Accuracy, the increase underscores the critical added value of polarimetry for enhancing the discrimination of hydrometeor types. Although the accuracy improvement is modest, polarimetric variables provide physically meaningful information on hydrometeor microphysics that cannot be captured by temperature alone. Given that many operational radars have polarimetric capabilities, including these variables leverages data already available and adds value beyond the specific accuracy gain observed in this experiment. Furthermore, the 74% Accuracy achieved by the RF_AROME configuration is consistent with recent studies showing that tree-based ensemble methods outperform traditional schemes in precipitation type classification [64].

Despite these improvements, some limitations persist. Agreement between the two systems remains strong primarily at the broader level of liquid versus solid precipitation. The mixed-phase category remains problematic, likely due to both the intrinsic complexity

of defining mixed precipitation and the unavoidable mismatch in spatial and temporal resolution among instruments. An additional contributing factor is that the disdrometer classifies mixed-phase precipitation very infrequently, since this phase occurs within a relatively narrow vertical region, the melting layer, which typically spans only a few hundred meters in depth [6]. Because the disdrometer measures at a very small sampling volume compared to the MRR or the X-band radar, it is much less unlikely to sample mixed hydrometeors within this limited vertical interval where mixed precipitation is present. Therefore, the observed overestimation by the radar does not necessarily indicate an error; rather, it reflects the inherent difference in sampling geometry.

In addition, the temporal resolution of the temperature data represents a further limitation. Both AWS observations and AROME model outputs are available at hourly resolution and are linearly interpolated to match the 5 min radar observations. This approximation is unable to capture abrupt temperature changes occurring at short time scales, particularly during frontal passages or rapidly evolving weather conditions. Such limitations are especially critical near the melting layer and are likely to affect the classification of mixed-phase precipitation.

Moreover, this study underscores the critical role of accurate vertical temperature profiles for operational radar-based precipitation phase classification. While obtaining such profiles can be challenging, especially in complex terrain, several viable strategies exist. These include the use of high-resolution numerical weather prediction models (e.g., AROME), in-situ AWS measurements at multiple elevations, and extrapolation methods based on the MRR-derived melting level height. Some studies have explored the possibility of having a vertical profile combining a microwave radiometer (MWR) with AROME forecast, with good results [65], or different methods to determine the melting level, such as [7,66,67]. These approaches have been successfully applied in previous research and represent practical solutions for operational settings.

A critical aspect that must be considered when interpreting the results of this study is the strong dependence on the vertical temperature profile. The analysis has been conducted exclusively over a single study area, and therefore the results cannot be directly generalized to other regions. Atmospheric and orographic characteristics vary significantly from one region to another, which can strongly influence the shape, variability, and representativeness of temperature profiles. Consequently, the optimal method for deriving the vertical temperature profile may differ across regions. For this reason, a regional validation should be performed to identify the most suitable temperature-profile estimation approach before applying the proposed classification framework in other locations.

Future studies should therefore focus on improving the representation of vertical temperature profiles in complex orography, with particular emphasis on enhancing their temporal and spatial resolution. This could involve the integration of additional in-situ observations at multiple elevations, the use of complementary remote sensing instruments, or hybrid approaches combining numerical weather prediction outputs with observational constraints. Improving the accuracy and continuity of temperature profiles is expected to significantly enhance the robustness and generalization capability of radar-based hydrometeor classification methods.

5. Conclusions

This study analyzed and compared different hydrometeor classification techniques derived from three instruments: a ground-based disdrometer, a Micro Rain Radar (MRR), and an operational X-band polarimetric radar. A unified methodology was developed and implemented to enable comparisons, despite the inherent differences in spatial and temporal sampling resolution among the instruments.

The principal findings and conclusions drawn from the analysis are summarized below:

- The MRR scheme, which provides hydrometeor classification along the vertical profile, was evaluated using three disdrometers. The results demonstrated a good level of agreement in classification, achieving an Accuracy of 86%. This high confidence level allowed the MRR classification to be reliably used as the primary reference for the subsequent evaluation of the XRAD classification schemes.
- The comprehensive analysis of all XRAD schemes clearly indicates the paramount importance of incorporating a temperature profile for accurate phase classification. Without a temperature profile, the Accuracy of the model is around 51%, while including the temperature profile increases it up to 74%. Three distinct approaches—based on AWS in-situ measurements, MRR observations, and AROME model forecasts—were used satisfactorily to derive the vertical temperature profiles.
- A machine learning random forest method (RF_AROME) was developed to classify hydrometeors from XRAD data. This scheme demonstrated superior performance, achieving better results than all other schemes across all evaluated verification scores, with an Accuracy of 78%.

In summary, the results highlight the value of using vertical profiling radars. The MRR not only enables reliable hydrometeor classification but also provides valuable insight into precipitation processes occurring directly below the XRAD beam. Furthermore, the results unequivocally demonstrate the capability of polarimetric radar to accurately classify precipitation type, provided a reliable temperature profile is included as an additional input variable.

Author Contributions: Conceptualization, B.B., F.P. and J.B.; methodology, B.B., F.P. and J.B.; formal analysis, F.P.; investigation, B.B., F.P., and J.B.; data curation, F.P.; writing—original draft preparation, F.P. and J.B.; writing—review and editing, A.G.-B., B.B., E.B., F.J.R., F.P., M.V., and M.U.; visualization, F.P.; project administration, J.B.; funding acquisition, J.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partly funded by the projects ARTEMIS (PID2021-124253OB-I00 MINECO/FEDER), the Water Research Institute (IdRA) of the University of Barcelona, Fundació Montcelimar, and AGAUR grant 2022 FISDU 00365.

Data Availability Statement: Data from IGE MRR, disdrometers, and AWS used in this study are available upon request to IGE center. Data from Météo-France AWS and X-Band radar are available on the Météo-France website (<https://donneespubliques.meteofrance.fr/> and <https://radarsmf.aeris-data.fr/en/home-page/>).

Acknowledgments: We thank Météo-France for providing the X-band radar weather data.

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

Abbreviations

The following abbreviations are used in this manuscript:

ASL	Above sea level
AWS	Automatic weather station
GSS	Gerrity skill score
HSS	Heidke skill score
MRR	Micro Rain Radar
XRAD	X-band polarimetric weather radar
PPI	Plan Position Indicator
RHI	Range Height Indicator
CSU	Colorado State University

IGE Institute of Environmental Geosciences
HCS Hydrometeor classification Summer
HCW Hydrometeor classification Winter

Appendix A

Table A1. Summary of the predictor set, target variable, pre-processing steps, QC thresholds, and hyperparameters used for the Random Forest classifier.

Aspect	Description
Target	MRR Hydrometeor phase classification (Rain, Mixed, Snow)
Predictors	XRAD values (Z_e , ρ_{HV} , ZDR, Kdp) and Temperature
Preprocessing	Aggregation to 5 min intervals (matching XRAD resolution); removal of missing values XRAD: Exclude minutes with $Z_e < 0$ dBZ.
QC thresholds	MRR: Precipitation considered only if $Z_e > 0$ dBZ for ≥ 3 consecutive minutes at ≥ 10 distinct vertical levels; vertical aggregation according to XRAD beams. Disdrometer: valid precipitation requires ≥ 3 consecutive minutes with > 50 detected particles per minute.
RF Hyperparameters	$n_estimators = 100$; $criterion = Gini$; $max_depth = 5$; $min_samples_split = 15$; $min_samples_leaf = 15$; $min_impurity_decrease = 0.001$; $class_weight = 'balanced'$

Appendix B

Table A2 summarizes the input data and specific name of each method used for hydrometeor classification. In total, 12 classification methods were applied: one for the disdrometer, one for the MRR, and ten for the XRAD. The XRAD methods include three main algorithms, each with sub-variants depending on the source of the temperature profile used. All classification methods were evaluated using the verification metrics described in Section 2.3.3 to enable a consistent comparison of their performance.

Table A2. Description of the different methodologies used for precipitation classification.

Input Data	Method	Characteristics
Disdrometer	Disdrometer	Disdrometer manufacturer software.
MRR	MRR	Precipitation type classification using MRR data and the RaProM program.
XRAD	HCW_NTemp	CSU classification using the detection of the freezing level from PPI or RHI scans, without a temperature profile.
	HCS_NTemp	CSU classification at a single pixel, without temperature information.
	HCS_AWS	CSU classification at a single pixel using temperature data from AWS.
	HCS_FL	CSU classification at a single pixel using temperature from the freezing level detection obtained by the MRR classification algorithm and surface temperature AWS.
	HCS_AROME	CSU classification at a single pixel using temperature from the AROME model.
	RF_Temp	Random forest classification at single pixel trained only with temperature data from AROME. Polarimetric variables are not used; Z_e is applied as a mask to determine precipitation occurrence.

Table A2. Cont.

Input Data	Method	Characteristics
	RF_NoTemp	Random forest classification at single pixel trained without temperature information.
	RF_AWS	Random forest classification at single pixel trained with XRAD variables and temperature from AWS.
	RF_FL	Random forest classification at single pixel trained with XRAD variables and temperature from the freezing level detection obtained by the MRR classification algorithm and surface temperature AWS.
	RF_AROME	Random forest classification at single pixel trained with XRAD variables and temperature profiles from AROME.

References

- Wrzesien, M.L.; Pavelsky, T.M.; Sobolowski, S.P.; Huning, L.S.; Cohen, J.S.; Herman, J.D. Tracking the Impacts of Precipitation Phase Changes Through the Hydrologic Cycle in Snowy Regions: From Precipitation to Reservoir Storage. *Front. Earth Sci.* **2022**, *10*, 995874. [\[CrossRef\]](#)
- Li, X.; Che, T.; Tang, Y.; Duan, H.; Wang, G.; Zhang, X.; Yang, C.; Wu, J.; Zhang, Y.; Li, L. The Shifts of Precipitation Phases and Their Impacts. *Sci. China Earth Sci.* **2025**, *68*, 425–443. [\[CrossRef\]](#)
- Roe, G.H.; Montgomery, D.R.; Hallet, B. Orographic Precipitation and the Relief of Mountain Ranges. *J. Geophys. Res.* **2003**, *108*, 2001JB001521. [\[CrossRef\]](#)
- Morán-Tejeda, E.; López-Moreno, J.I.; Sanmiguel-Vallado, A. Changes in Climate, Snow and Water Resources in the Spanish Pyrenees: Observations and Projections in a Warming Climate. In *High Mountain Conservation in a Changing World*; Catalan, J., Ninot, J.M., Aniz, M.M., Eds.; Advances in Global Change Research; Springer International Publishing: Cham, Switzerland, 2017; Volume 62, pp. 305–323.
- Spandre, P.; François, H.; Verfaillie, D.; Pons, M.; Vernay, M.; Lafaysse, M.; George, E.; Morin, S. Winter Tourism under Climate Change in the Pyrenees and the French Alps: Relevance of Snowmaking as a Technical Adaptation. *Cryosphere* **2019**, *13*, 1325–1347. [\[CrossRef\]](#)
- Van Den Heuvel, F.; Gabella, M.; Germann, U.; Berne, A. Characterisation of the Melting Layer Variability in an Alpine Valley Based on Polarimetric X-Band Radar Scans. *Atmos. Meas. Tech.* **2018**, *11*, 5181–5198. [\[CrossRef\]](#)
- Khanal, A.K.; Delrieu, G.; Cazenave, F.; Boudevillain, B. Radar Remote Sensing of Precipitation in High Mountains: Detection and Characterization of Melting Layer in the Grenoble Valley, French Alps. *Atmosphere* **2019**, *10*, 784. [\[CrossRef\]](#)
- Leroux, N.R.; Vionnet, V.; Thériault, J.M. Performance of Precipitation Phase Partitioning Methods and Their Impact on Snowpack Evolution in a Humid Continental Climate. *Hydrol. Processes* **2023**, *37*, e15028. [\[CrossRef\]](#)
- Marks, D.; Winstral, A.; Reba, M.; Pomeroy, J.; Kumar, M. An Evaluation of Methods for Determining During-Storm Precipitation Phase and the Rain/Snow Transition Elevation at the Surface in a Mountain Basin. *Adv. Water Resour.* **2013**, *55*, 98–110. [\[CrossRef\]](#)
- Dolan, B.; Rutledge, S.A. A Theory-Based Hydrometeor Identification Algorithm for X-Band Polarimetric Radars. *J. Atmos. Ocean. Technol.* **2009**, *26*, 2071–2088. [\[CrossRef\]](#)
- Thompson, E.J.; Rutledge, S.A.; Dolan, B.; Chandrasekar, V.; Cheong, B.L. A Dual-Polarization Radar Hydrometeor Classification Algorithm for Winter Precipitation. *J. Atmos. Ocean. Technol.* **2014**, *31*, 1457–1481. [\[CrossRef\]](#)
- Besic, N.; Figueras, I.; Ventura, J.; Grazioli, J.; Gabella, M.; Germann, U.; Berne, A. Hydrometeor Classification Through Statistical Clustering of Polarimetric Radar Measurements: A Semi-Supervised Approach. *Atmos. Meas. Tech.* **2016**, *9*, 4425–4445. [\[CrossRef\]](#)
- Seo, B.-C. A Data-Driven Approach for Winter Precipitation Classification Using Weather Radar and NWP Data. *Atmosphere* **2020**, *11*, 701. [\[CrossRef\]](#)
- 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. [\[CrossRef\]](#)
- Garcia-Benadí, A.; Bech, J.; Gonzalez, S.; Udina, M.; Codina, B.; Georgis, J.F. Precipitation Type Classification of Micro Rain Radar Data Using an Improved Doppler Spectral Processing Methodology. *Remote Sens.* **2020**, *12*, 4113. [\[CrossRef\]](#)
- Garcia-Benadí, A.; Bech, J.; Gonzalez, S.; Udina, M.; Codina, B. A New Methodology to Characterise the Radar Bright Band Using Doppler Spectral Moments from Vertically Pointing Radar Observations. *Remote Sens.* **2021**, *13*, 4323. [\[CrossRef\]](#)
- Zhou, L.; Zhou, Y.; Lu, C.; Lü, J.; Fu, Z. Comparative Study on Microphysical Characteristics of Two Icing and Freezing Weather Events over Central China in February 2024. *Atmos. Res.* **2026**, *331*, 108688. [\[CrossRef\]](#)

18. Faure, D.; Delrieu, G.; Gaussiat, N. Impact of the Altitudinal Gradients of Precipitation on the Radar QPE Bias in the French Alps. *Atmosphere* **2019**, *10*, 306. [\[CrossRef\]](#)
19. Bacer, S.; Beaumet, J.; Ménégot, M.; Gallée, H.; Le Bouëdec, E.; Staquet, C. Impact of Climate Change on Persistent Cold-Air Pools in an Alpine Valley During the 21st Century. *Weather. Clim. Dyn.* **2024**, *5*, 211–229. [\[CrossRef\]](#)
20. Peel, M.C.; Finlayson, B.L.; McMahon, T.A. Updated World Map of the Köppen-Geiger Climate Classification. *Hydrol. Earth Syst. Sci.* **2007**, *11*, 1633–1644. [\[CrossRef\]](#)
21. Figueras, I.; Ventura, J.; Tabary, P. The New French Operational Polarimetric Radar Rainfall Rate Product. *J. Appl. Meteorol. Climatol.* **2013**, *52*, 1817–1835. [\[CrossRef\]](#)
22. Doviak, R.J.; Zrnic, D.S. *Doppler Radar and Weather Observations*; Dover Publications, Inc.: Garden, NY, USA, 1993.
23. Delrieu, G.; Andrieu, H.; Creutin, J.D. Quantification of Path-Integrated Attenuation for X- and C-Band Weather Radar Systems Operating in Mediterranean Heavy Rainfall. *J. Appl. Meteorol.* **2000**, *39*, 840–850. [\[CrossRef\]](#)
24. Helmus, J.J.; Collis, S.M. The Python ARM Radar Toolkit (Py-ART), a Library for Working with Weather Radar Data in the Python Programming Language. *J. Open Res. Softw.* **2016**, *4*, 25. [\[CrossRef\]](#)
25. Bringi, V.N.; Chandrasekar, V. *Polarimetric Doppler Weather Radar: Principles and Applications*, 1st ed.; Cambridge University Press: Cambridge, UK, 2001.
26. Kumjian, M. Principles and Applications of Dual-Polarization Weather Radar. Part I: Description of the Polarimetric Radar Variables. *J. Oper. Meteorol.* **2013**, *1*, 226–242. [\[CrossRef\]](#)
27. Kumjian, M. Principles and Applications of Dual-Polarization Weather Radar. Part II: Warm- and Cold-Season Applications. *J. Oper. Meteorol.* **2013**, *1*, 243–264. [\[CrossRef\]](#)
28. Vulpiani, G.; Montopoli, M.; Passeri, L.D.; Gioia, A.G.; Giordano, P.; Marzano, F.S. On the Use of Dual-Polarized C-Band Radar for Operational Rainfall Retrieval in Mountainous Areas. *J. Appl. Meteorol. Climatol.* **2012**, *51*, 405–425. [\[CrossRef\]](#)
29. Löffler-Mang, M.; Kunz, M.; Schmid, W. On the Performance of a Low-Cost K-Band Doppler Radar for Quantitative Rain Measurements. *J. Atmos. Ocean. Technol.* **1999**, *16*, 379–387. [\[CrossRef\]](#)
30. Peters, G.; Bernd, F.; Andersson, T. Rain Observations with a Vertically Looking Micro Rain Radar (MRR). *Boreal. Environ. Res.* **2002**, *7*, 353–362.
31. Micro-Rain-Radar MRR-2 and MRR-PRO. A Tutorial, Metek. Available online: https://metek.de/wp-content/uploads/2016/12/2018-0206-MRR_tutorial.pdf (accessed on 10 March 2025).
32. Méndez, B.; Saenz, E.; Pires, Ó.; Cantero, E.; Bech, J.; Polls, F.; Peinó, E.; Udina, M.; Garcia-Benadí, A. Experimental Campaign for the Characterization of Precipitation in a Complex Terrain Site Using High Resolution Observations. *J. Phys. Conf. Ser.* **2024**, *2767*, 042016. [\[CrossRef\]](#)
33. Reinoso-Rondinel, R.; Schleiss, M. Quantitative Evaluation of Polarimetric Estimates from Scanning Weather Radars Using a Vertically Pointing Micro Rain Radar. *J. Atmos. Ocean. Technol.* **2021**, *38*, 481–499. [\[CrossRef\]](#)
34. Blahak, U. An Approximation to the Effective Beam Weighting Function for Scanning Meteorological Radars with an Axisymmetric Antenna Pattern. *J. Atmos. Ocean. Technol.* **2008**, *25*, 1182–1196. [\[CrossRef\]](#)
35. Löffler-Mang, M.; Joss, J. An Optical Disdrometer for Measuring Size and Velocity of Hydrometeors. *J. Atmos. Ocean. Technol.* **2000**, *17*, 130–139. [\[CrossRef\]](#)
36. Tokay, A.; Wolff, D.B.; Petersen, W.A. Evaluation of the New Version of the Laser-Optical Disdrometer, OTT Parsivel2. *J. Atmos. Ocean. Technol.* **2014**, *31*, 1276–1288. [\[CrossRef\]](#)
37. Casanovas, C.; Salio, P.; Galligani, V.; Dolan, B.; Nesbitt, S.W. Drop Size Distribution Variability in Central Argentina During Relampago-Cacti. *Remote Sens.* **2021**, *13*, 2026. [\[CrossRef\]](#)
38. Dolan, B.; Fuchs, B.; Rutledge, S.A.; Barnes, E.A.; Thompson, E.J. Primary Modes of Global Drop Size Distributions. *J. Atmos. Sci.* **2018**, *75*, 1453–1476. [\[CrossRef\]](#)
39. Thompson, E.J.; Rutledge, S.A.; Dolan, B.; Thurai, M. Drop Size Distributions and Radar Observations of Convective and Stratiform Rain over the Equatorial Indian and West Pacific Oceans. *J. Atmos. Sci.* **2015**, *72*, 4091–4125. [\[CrossRef\]](#)
40. Guyot, A.; Pudashine, J.; Protat, A.; Uijlenhoet, R.; Pauwels, V.R.N.; Seed, A.; Walker, J.P. Effect of Disdrometer Type on Rain Drop Size Distribution Characterisation: A New Dataset for South-Eastern Australia. *Hydrol. Earth Syst. Sci.* **2019**, *23*, 4737–4761. [\[CrossRef\]](#)
41. Loh, J.L.; Lee, D.I.; Kang, M.Y.; You, C.H. Classification of Rainfall Types Using Parsivel Disdrometer and S-Band Polarimetric Radar in Central Korea. *Remote Sens.* **2020**, *12*, 642. [\[CrossRef\]](#)
42. Ulbrich, C.W.; Atlas, D. Rainfall Microphysics and Radar Properties: Analysis Methods for Drop Size Spectra. *J. Appl. Meteorol.* **1998**, *37*, 912–923. [\[CrossRef\]](#)
43. Vivekanandan, J.; Zhang, G.; Brandes, E. Polarimetric Radar Estimators Based on a Constrained Gamma Drop Size Distribution Model. *J. Appl. Meteorol.* **2004**, *43*, 217–230. [\[CrossRef\]](#)
44. Gunn, R.; Kinzer, G.D. The Terminal Velocity of Fall for Water Droplets in Stagnant Air. *J. Meteorol.* **1949**, *6*, 243–248. [\[CrossRef\]](#)

45. Serio, M.A.; Carollo, F.G.; Ferro, V. Raindrop Size Distribution and Terminal Velocity for Rainfall Erosivity Studies. A Review. *J. Hydrol.* **2019**, *576*, 210–228. [\[CrossRef\]](#)
46. Ebert, E.E. Fuzzy Verification of High-Resolution Gridded Forecasts: A Review and Proposed Framework. *Meteorol. Appl.* **2008**, *15*, 51–64. [\[CrossRef\]](#)
47. Seity, Y.; Brousseau, P.; Malardel, S.; Hello, G.; Bénard, P.; Bouttier, F.; Lac, C.; Masson, V. The AROME-France Convective-Scale Operational Model. *Mon. Weather. Rev.* **2011**, *139*, 976–991. [\[CrossRef\]](#)
48. Gleeson, E.; Kurzeneva, E.; De Rooy, W.; Rontu, L.; Martín Pérez, D.; Clancy, C.; Ivarsson, K.-I.; Engdahl, B.J.; Tijm, S.; Nielsen, K.P.; et al. The Cycle 46 Configuration of the HARMONIE-AROME Forecast Model. *Meteorology* **2024**, *3*, 354–390. [\[CrossRef\]](#)
49. Bédard-Therrien, A.; Anctil, F.; Thériault, J.M.; Chalifour, O.; Payette, F.; Vidal, A.; Nadeau, D.F. Leveraging a Radar-Based Disdrometer Network to Develop a Probabilistic Precipitation Phase Model in Eastern Canada. *Hydrol. Earth Syst. Sci.* **2025**, *29*, 1135–1158. [\[CrossRef\]](#)
50. Lang, T.; Dolan, B.; Guy, N.; Gerlach, C.A.M.; Hardin, J.; Raut, B. *CSU-Radarmet/CSU_RadarTools: CSU_RadarTools v1.4 2024*; Zenodo: Genève, Switzerland, 2024.
51. 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. Climatol.* **2013**, *52*, 2162–2186. [\[CrossRef\]](#)
52. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [\[CrossRef\]](#)
53. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-Learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
54. Largeron, Y.; Staquet, C. Persistent Inversion Dynamics and Wintertime PM10 Air Pollution in Alpine Valleys. *Atmos. Environ.* **2016**, *135*, 92–108. [\[CrossRef\]](#)
55. Gerosa, G.A.; Finco, A.; Giovannini, L.; Zardi, D.; Marzuoli, R. Comparing Airborne and Along-Sidewall Ground-Based Measurements to Retrieve Thermal and PM10 Concentration Profiles in Two Alpine Valleys. *J. Appl. Meteorol. Climatol.* **2024**, *63*, 1327–1341. [\[CrossRef\]](#)
56. Jolliffe, I.T.; Stephenson, D. *Forecast Verification: A Practitioner's Guide in Atmospheric Science*; J. Wiley: Chichester, UK, 2003.
57. Wilks, D.S. *Statistical Methods in the Atmospheric Sciences*, 2nd ed.; International Geophysics Series; Elsevier: Amsterdam, The Netherlands; Paris, France, 2006.
58. Sokolova, M.; Lapalme, G. A Systematic Analysis of Performance Measures for Classification Tasks. *Inf. Process. Manag.* **2009**, *45*, 427–437. [\[CrossRef\]](#)
59. Heidke, P. Berechnung Des Erfolges Und Der Güte Der Windstärkevorhersagen Im Sturmwarnungsdienst. *Geogr. Ann.* **1926**, *8*, 301–349. [\[CrossRef\]](#)
60. Gerrity, J.P. A Note on Gandin and Murphy's Equitable Skill Score. *Mon. Weather Rev.* **1992**, *120*, 2709–2712. [\[CrossRef\]](#)
61. Brooks, H.E.; Flora, M.L.; Baldwin, M.E. A Rose by Any Other Name: On Basic Scores from the 2×2 Table and the Plethora of Names Attached to Them. *Artif. Intell. Earth Syst.* **2024**, *3*, e230104. [\[CrossRef\]](#)
62. Frech, M.; Hagen, M.; Mammen, T. Monitoring the Absolute Calibration of a Polarimetric Weather Radar. *J. Atmos. Ocean. Technol.* **2017**, *34*, 599–615. [\[CrossRef\]](#)
63. 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. [\[CrossRef\]](#)
64. 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. [\[CrossRef\]](#)
65. Martinet, P.; Cimini, D.; De Angelis, F.; Canut, G.; Unger, V.; Guillot, R.; Tzanos, D.; Paci, A. Combining Ground-Based Microwave Radiometer and the AROME Convective Scale Model Through 1DVAR Retrievals in Complex Terrain: An Alpine Valley Case Study. *Atmos. Meas. Tech.* **2017**, *10*, 3385–3402. [\[CrossRef\]](#)
66. Brast, M.; Markmann, P. Detecting the Melting Layer with a Micro Rain Radar Using a Neural Network Approach. *Atmos. Meas. Tech.* **2020**, *13*, 6645–6656. [\[CrossRef\]](#)
67. Ryzhkov, A.; Krause, J. New Polarimetric Radar Algorithm for Melting-Layer Detection and Determination of Its Height. *J. Atmos. Ocean. Technol.* **2022**, *39*, 529–543. [\[CrossRef\]](#)

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.