Hydrometeor classification from two-dimensional video disdrometer data

J. Grazioli[1], **D. Tuia**[2], **S. Monhart**[3], **M. Schneebeli**[3], **T. Raupach**[1], **and A. Berne**[1]

[1]Environmental Remote Sensing Laboratory (LTE), École Polytechnique Fédérale de Lausanne (EPFL),
Lausanne, Switzerland
[2]Laboratory of Geographic Information Systems (LASIG), École Polytechnique Fédérale de Lausanne (EPFL),
Lausanne, Switzerland
[3]Federal Office of Meteorology and Climatology, Meteo Swiss, Locarno-Monti, Switzerland

Correspondence to: A. Berne (alexis.berne@epfl.ch)

Abstract. The first hydrometeor classification technique based on two-dimensional video disdrometer (2DVD) data is presented. The method provides an estimate of the dominant hydrometeor type falling over time intervals of 60 s during precipitation, using the statistical behavior of a set of particle descriptors as input, calculated for each particle image. The employed supervised algorithm is a support vector machine (SVM), trained over 60 s precipitation time steps labeled by visual inspection. In this way, eight dominant hydrometeor classes can be discriminated. The algorithm achieved high classification performances, with median overall accuracies (Cohen's $\mathcal{K}$) of 90 % (0.88), and with accuracies higher than 84 % for each hydrometeor class.

1 Introduction

The two-dimensional video disdrometer (Kruger and Krajewski, 2002), 2DVD hereafter, significantly improves the capability of ground observations to describe the microphysics and microstructure of precipitation both in the solid and the liquid phase. The system, based on simultaneous observations of falling objects with two orthogonally oriented cameras, has been used to characterize the relationships linking raindrop shape, size, and terminal velocity (e.g., Thurai and Bringi, 2005; Thurai et al., 2009). It has also been employed to validate weather radar rainfall estimates (Schuur et al., 2001; Thurai et al., 2008; Cao et al., 2008; Zhang et al., 2008). Regarding snowfall, the 2DVD has been used to

derive the statistical properties of particle size distributions of winter storms (Brandes et al., 2007), to improve the radar retrieval of equivalent liquid precipitation (Huang et al., 2010), and to simulate radar observations from measured snowfall microstructure (Zhang et al., 2011).

In the present work we employ 2DVD measurements for the classification of hydrometeors, with a special focus on ice-phase precipitation. The expression "hydrometeor classification" refers to techniques that aim to retrieve qualitative information about the dominant hydrometeor type that characterizes the precipitation. Such information can then be used for risk assessment (hazardous hydrometeor identification, like hail), for parametrization and validation of numerical weather prediction (NWP) models (e.g., Xue et al., 2000), or as support for microphysical investigations (e.g., Houze, 1993; Schneebeli et al., 2013). Today, hydrometeor classification techniques are implemented for different types of measurements. Typical examples in remote sensing are algorithms designed for ground-based polarimetric weather radars (Straka et al., 2000; Dolan and Rutledge, 2009; Chandrasekar et al., 2013), or for airborne radars and lidars observing ice-phase clouds (e.g., Shupe, 2007; Delanoe et al., 2013). These sensors enable the sampling of large domains at high resolution on a short timescale, but their classification retrievals are indirect, constrained by numerical simulations, and difficult to validate extensively. In contrast, airborne particle probe imagers (e.g., Feind, 2008) allow for direct classification along aircraft flight paths but only (given the high cost of these platforms) during intensive

measurement campaigns. Ground-based instruments sample precipitation directly on site (although on small sampling volumes), and could be used to classify hydrometeors, thus becoming a point reference for remote sensing retrievals. Only a few research works have been devoted to the implementation of classification schemes for such instruments, and their focus was mostly on mixed-phase precipitation (Yuter et al., 2006) or the exploration of the potential synergy between multiple sensors (Marzano et al., 2010). Some commercial disdrometers (e.g., PARSIVEL; Löffler-Mang and Joss, 2000), originally designed for rainfall studies, provide a basic estimation of the precipitation type associated with each measurement by making assumptions on fall velocity and equivalent rainfall intensity.

In this context the information provided by the 2DVD is of particular interest because a pair of two-dimensional views, together with fall velocity, is provided for each particle. Such features alone allow expert users to interpret the images and visually recognize in them specific hydrometeor types (e.g., Zhang et al., 2011). This suggests that automatic classification methods, based on training over visually interpreted (labeled) episodes, may be well suited to perform hydrometeor classification. Supervised classification algorithms, such as the support vector machine, (SVM; Boser et al., 1992), are used today to perform similar kinds of tasks. For example, such techniques have been used in land cover classification (Camps-Valls and Bruzzone, 2005), wind power forecasts (Foresti et al., 2011; Zeng and Qiao, 2011), and weather prediction (Sullivan, 2009). The SVM is a linear and binary supervised classifier that finds the optimal separations between observations belonging to different classes. These observations are defined by a set of numerical features, and the optimal separation is learned from a training set in which the association between input observation and output class is known. The SVM is able to handle high-dimensional inputs, is less prone to over-fitting issues than other supervised methods (Camps-Valls and Bruzzone, 2009), and has been shown to perform relatively well on the prediction of weather types (e.g., Elmore, 2010). Furthermore, the SVM allows for the retrieval of the most relevant input features driving the classification, and can rank them in order of importance, with the implementation of multiple kernel learning (SVM-MKL) techniques (Rakotomamonjy et al., 2008; Tuia et al., 2010).

In this paper we train an SVM model on 2DVD data in order to classify eight hydrometeor classes of the dominant type of precipitation during time intervals of length Δt. Aggregation over time intervals is conducted to reduce the computational cost, which may be excessive if each particle is individually considered. A relatively short Δt of 60 s is chosen to minimize the effect of mixing of separate hydrometeor types. Individual 2DVD images are summarized over Δt with a high-dimensional set of numerical features, constituting the necessary input for the SVM classifier. Data collected in the Swiss Alps; in the French Jura; and in the southern part of Ontario, Canada, are used to train and validate the model.
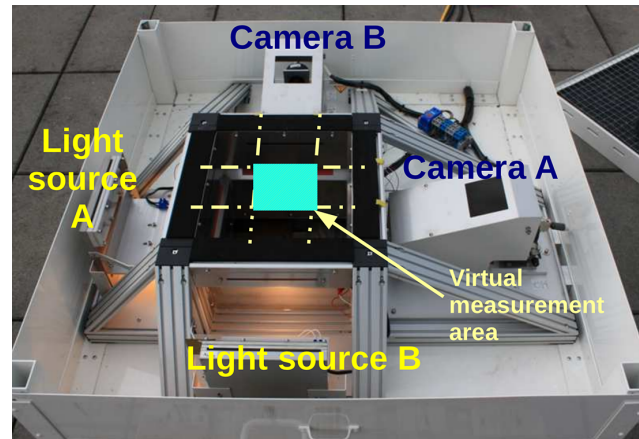

Figure 1. 2DVD measurement setup.

The manuscript is structured as follows. Section 2 describes the experimental setup and the basic 2DVD data. Section 3 presents the hydrometeor classification model. Section 4 presents the main results and their quality assessment, while Sect. 5 provides examples of the outputs of the hydrometeor classification. Section 6 concludes the paper and lists some future perspectives.

2 Data set description

2.1 Experiment locations

The 2DVD data employed in the experiments were collected during three distinct field campaigns, between September 2009 and March 2013. The first campaign took place from September 2009 until June 2011 in Davos, Switzerland: the 2DVD was deployed in the Swiss Alps, at an altitude of about 2500 m a.s.l. Data for a total of 1700 h of precipitation in liquid, mixed, and solid phase were collected during this time frame. The second campaign took place in Remoray, France, from December 2012 until March 2013, at an altitude of about 920 m, in the context of an experiment focused on melting hydrometeors. A total of 270 h of precipitation in solid, liquid, and mixed phase was collected in this experiment. The third complementary campaign includes about 200 h of data (mainly solid precipitation) collected by three 2DVD instruments between December 2011 and March 2012 in the framework of the Global Precipitation Measurement mission (GPM, http://pmm.nasa.gov/precipitation-measurement-missions), in the Cold-season Precipitation Experiment (GCPEx) that took place in Ontario, Canada.

2.2 2DVD instrument and data pre-processing

The 2DVD working principle is extensively described in Kruger and Krajewski (2002). Here we briefly summarize the
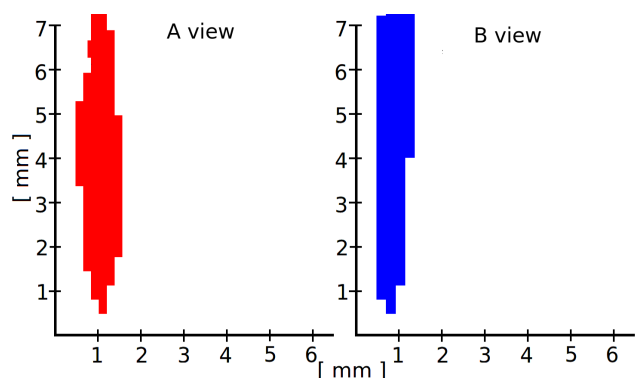
Figure 2. Example of A–B views of a non-realistic particle that needs to be filtered.

most relevant features of the instrument. Figure 1 illustrates the 2DVD measurement principle (see Fig. 3 of Kruger and Krajewski, 2002, for more details).

Two orthogonal light sources coupled with two (A and B) line-scanning cameras generate two stacked measurement planes of about 11 cm × 11 cm. The planes are vertically separated by a distance of around 6.2–7 mm (the exact value is determined by mechanical calibration). The cameras capture the falling particles at a resolution of 512 pixels (0.2 mm) at 34 kHz, and the vertical distance between the measurement areas of cameras A and B enables the measurement of fall velocity.

The raw images need to be processed before being employed. This involves the filtering of unreasonable measurements, as well as the rematching of the measurements taken from camera A and B in order to ensure that both images actually refer to the same particle. Filtering and rematching of 2DVD images is based on the work of Hanesch (1999) and Huang et al. (2010). We followed their methods, but with a noteworthy modification. Those studies, which were interested in snowfall only, restricted the maximum fall velocity to 4 and 6 m s^{-1}, respectively. We increased this upper boundary to 14 m s^{-1}, large enough to include, with sufficient margins, the range of variation found in rain (e.g., Beard, 1976) and large graupel (List and Schemena, 1971).

Despite this filtering, some non-realistic particles can still be observed in the output. These particles appear as large objects, vertically oriented and elongated, as shown in Fig. 2. Because of these peculiarities, they are easily identified and excluded from the analysis presented in this paper. The exact nature of these artifacts is unknown, but their vertical orientation and dimension suggest that they may be associated with small-scale wind effects, melting, or dripping, causing some particles to reside for an anomalous amount of time in the measurement areas of the two cameras. The proportion of rejected particles is on average 3 %, and it ranges between 0.5 and 13 % per day. A few precipitation events required higher rejection rates. They were excluded from the analysis presented in this work.

Table 1. List of descriptors chosen to describe the particles recorded. Descriptors 1 and 2 come from the combination of camera A and B; 3 to 6 describe particle size, and 7 to 13 particle shape.

	Symbol	Full name	Units
1	v	Fall velocity	[m s^{-1}]
2	D_e	Equivolumetric diameter	[mm]
3	$A_{A,B}$	Shaded area	[mm^2]
4	$P_{A,B}$	Shaded perimeter	[mm]
5	$T_{A,B}$	Particle thickness	[mm]
6	$W_{A,B}$	Particle width	[mm]
7	PF$_{A,B}$	Pixel fraction	[–]
8	FORM$_{A,B}$	Form index	[–]
9	SqP$_{A,B}$	Square pixel metric	[–]
10	FD$_{A,B}$	Fractal dimension	[–]
11	SI$_{A,B}$	Shape index	[–]
12	ELONG$_{A,B}$	Elongation	[–]
13	ROUND$_{A,B}$	Roundness	[–]

Two additional potential sources of uncertainty (whose magnitude is currently not known in snowfall) are the image distortion that can occur when the horizontal component of the falling velocity of the particles is significant, and the local winds generated by the geometry of the instrument. To date, image distortion can only be corrected in rain, and in particular only for raindrops that possess an axis of rotational symmetry (Schönhuber et al., 2008). In contrast, the winds induced by the instrument itself can lead to an underestimation of particles having lower density and dimension[1]. Further research, which is beyond the scope of this paper, is needed to develop correction schemes for snowfall measurements in order to compensate for these two potential issues.

2.3 From single particles to global features

Pairs of 2DVD A–B images are available for each particle falling in the measurement area. For the purpose of the present work, it is useful to summarize this large amount of information by choosing a set of relevant descriptors[2]. Then, the statistical distributions of these descriptors in a time step Δt are used as input information for the hydrometeor classification. The descriptors chosen in this work are listed in Table 1, and can be divided into three groups.

2.3.1 Joint descriptors

Two descriptors are obtained by combining the views of cameras A and B. They are particle falling velocity v [m s^{-1}], and

---

[1]This issue is more severe for the first generation of the 2DVD instrument (Nespor et al., 2000). All the data employed in the present study were collected with second- and third-generation 2DVDs.

[2]The particle descriptors are calculated in the present work from ".SNO" 2DVD data.
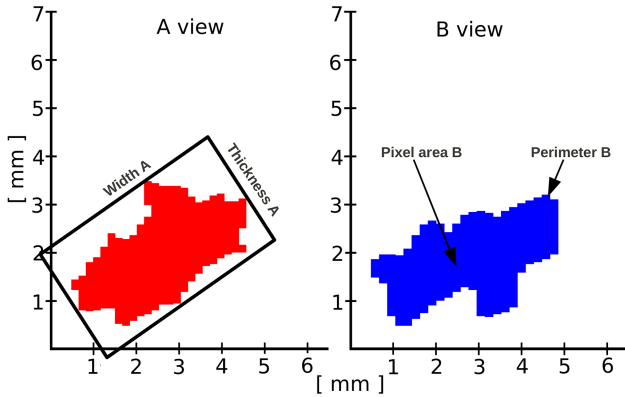
Figure 3. Examples of particles descriptors of 2DVD images (camera A and B). On camera A: width (W_A [mm]) and thickness (T_A [mm]) of the bounding box enclosing the particle. On camera B: particle apparent perimeter (P_A [mm]) and shaded area (A_A [mm^2]).

equivolumetric diameter D_e [mm]. D_e denotes the diameter of a sphere with the same volume as the falling particle.

This descriptor was originally developed for raindrops, for which volumes could be calculated accurately from the 2-D views. It can be extended to particles of any shape as a reference measure of particle size. In the present work, D_e is calculated according to the formulation of Huang et al. (2010).

2.3.2 Particle size

Other descriptors can be computed separately for camera A and camera B. Figure 3 illustrates some of them. The apparent shaded areas $A_{A,B}$ and perimeters $P_{A,B}$ are readily available from the 2DVD measurements, while thicknesses $T_{A,B}$ and widths $W_{A,B}$ of each particle are defined with respect to a bounding box around the particle (Fig. 3). v, D_e, A, P, T, and W together describe the particle bulk dimension and velocity.

2.3.3 Particle shape

Additional descriptors are computed to better characterize particle shape. They are dimensionless shape metrics commonly used in the analysis of land cover images for remote sensing (Jiao and Liu, 2012), adapted for use on 2DVD images:

$$\text{PF}_{A,B} = \frac{A_{A,B}}{A_{A,B}^r} \quad (0, 1], \tag{1}$$

$$\text{FORM}_{A,B} = \frac{4\pi A_{A,B}}{P_{A,B}^2} \quad (0, 1], \tag{2}$$

$$\text{SqP}_{A,B} = 1 - 4\frac{\sqrt{A_{A,B}}}{P_{A,B}} \quad [1 - 2/\sqrt{\pi}, 1), \tag{3}$$

$$\text{FD}_{A,B} = 2\frac{\ln(P_{A,B}/4)}{\ln(A_{A,B})} \quad [1, 2], \tag{4}$$

$$\text{SI}_{A,B} = \frac{P_{A,B}}{4\sqrt{A_{A,B}}} \quad [\sqrt{\pi}/2, +\infty], \tag{5}$$

$$\text{ELONG}_{A,B} = \frac{W_{A,B}}{T_{A,B}} \quad [1, +\infty], \tag{6}$$

$$\text{ROUND}_{A,B} = 4\frac{A_{A,B}}{\pi W_{A,B}^2} \quad (0, 1], \tag{7}$$

where $A_{A,B}^r$ [mm^2] is the area of the bounding box calculated for image A (or B). $\text{PF}_{A,B}$ is the pixel fraction and compares the shaded area with the area of the bounding box. $\text{PF}_{A,B}$ is an index of compactness, as is the roundness index ($\text{ROUND}_{A,B}$), which compares the shaded area with a circular approximation. $\text{FORM}_{A,B}$ and square pixel metric $\text{SqP}_{A,B}$ are shape complexity indices based on the area-to-perimeter ratio (they increase with decreasing complexity), while fractal dimension $\text{FD}_{A,B}$ and shape index $\text{SI}_{a,b}$ are indexes based on the perimeter-to-area ratio (they increase with increasing complexity). $\text{ELONG}_{A,B}$ evaluates the degree of elongation of the particles.

As introduced above, the feature vector used in the SVM model refers to the distribution of descriptors in a time step Δt. Let us consider a time step, Δt, during which N particles are recorded. We compute the mean, median, some quantiles (10, 25, 75, 90 %), and interquantiles (Q75–25, Q90–10) of each descriptor over the N particles available. Additionally, for descriptors 3 to 13 of Table 1, we compute the correlation coefficient between the measurements of camera A and B. This leads to a set of 203 features calculated per time step: 16 derived from camera combinations, 88 calculated separately for A and B (so 176 in total), and 11 correlation coefficients. We selected Δt to be 60 s, as a trade-off between representativeness and temporal resolution. Additionally, no statistics are computed if N is lower than 20 particles for a specific time step (Appendix A). The 88 features calculated separately for A and B have been verified to be consistent between each other, with biases generally lower than 10 %. This suggests that the information carried by a single camera is sufficient for these 88 features. Therefore we can define, for each valid time step, a final feature vector $\boldsymbol{x}$ containing 115 useful features by using only the 88 single features from one of the two cameras.
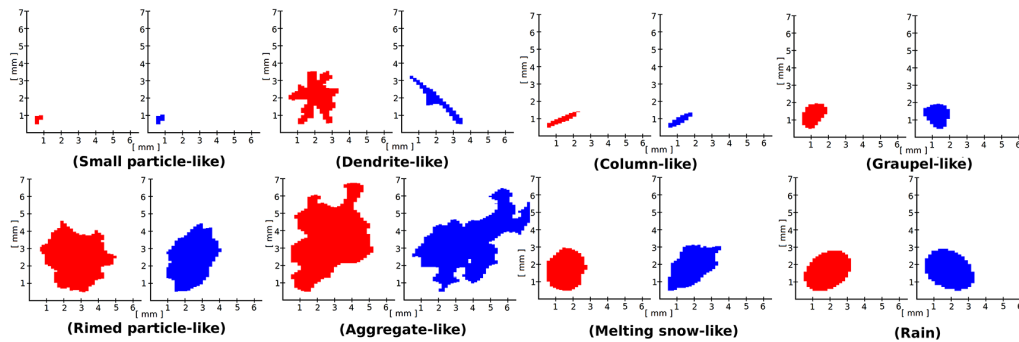
Figure 4. Examples of particle images (two camera views: A left, B right) belonging to time steps dominated by a particular hydrometeor class.

3 Hydrometeor classification

This section details the proposed supervised classification approach. First we define the hydrometeor classes, then we detail how a training set is obtained, and finally we present the classifier employed and its implementation to the available data set.

3.1 Hydrometeor classes and training set

The principle of supervised classification methods is to use a set of N_{train} labeled observations (or a training set) to train a classifier that will learn how to interpret new unlabeled observations. In our case, we need to assign the appropriate dominant hydrometeor type to a selected population of time steps of length Δt. The 2DVD offers the possibility to visualize the actual hydrometeor images, and the supervision was therefore conducted manually, according to the judgement of trained operators. Two operators independently interpreted the images by visualizing particle shapes, velocities, and taking into account the on-site environmental conditions (time of the year, temperature). Additionally, for the data collected in Davos, X-band radar observations over the region were available (e.g., Schneebeli et al., 2013), thus providing contextual information about the structure of the precipitation and, in stratiform cases, about the altitude of the melting layer.

The visualization and pre-interpretation of a wide range of time steps led to the selection of eight hydrometeor classes to describe the possible precipitation types in the available data set. Figure 4 shows an example of a typical particle belonging to each class. The classes are small-particle-like (SP), dendrite-like (D), column-like (C), graupel-like (G), rimed-particle-like (RIM), aggregate-like (AG), melting-snow-like (MS), and rain (R). The "-like" is added to emphasize that this approach identifies the dominant type of hydrometeor, which does not necessarily imply that (i) there is only one type of hydrometeor in the considered time step and (ii) that all hydrometeors exhibit pristine shape and geometry.

The definitions for some hydrometeor classes require clarification. SP time steps refer to particles falling during ice-phase precipitation that, given their size and the resolution of the instrument (0.2 mm), do not allow for proper visual shape recognition. Small aggregates, as well as single ice crystals, can be assumed to belong to this class. RIM is observed when riming processes smooth the shapes of the hydrometeors and increase their fall speed, while G time steps refer to fully developed graupel in which the original shape of the rimed crystal is no longer recognizable. MS is observed when the instrument records precipitation within the melting layer, and in these time steps raindrops, snowflakes, and smoother snowflakes with larger fall speed coexist in a mixed phase.

The creation of the training set involved the inspection of all the particles within each time step, in order to retrieve the dominant particle type and to provide the appropriate label. Particular attention was paid to select time steps that were as pure as possible, for the subsequent training of the classifier. The training set employed in the present work includes $N_{\text{train}} = 400$ time steps, each of them numerically characterized by the 115 components of the associated feature vector $\boldsymbol{x}$ defined in Sect. 2.3.

3.2 Classification method

In this section, we present the classifier used, the SVM. Then, we briefly detail an extension of the SVM that allows for the retrieval of the importance of each input feature (or group of features) in the model: SVM-MKL.

3.2.1 SVM

The support vector machine (Boser et al., 1992; Scholkopf and Smola, 2001), also known as the large margin classifier, is a linear classifier which finds the best linear separation between samples belonging to two classes. In our case, samples are time steps i of length Δt, represented by a vector $\boldsymbol{x}_i$ of $d = 115$ features, and the classes are the dominant types of hydrometeors, y_i. The model is trained on known couples $X_{\text{train}} = \{\boldsymbol{x}_i, y_i\}_{i=1}^{N_{\text{train}}}$, with $\boldsymbol{x}_i \in \mathbb{R}^d$ and $y_i \in [-1, 1]$. It must

generalize well on a set of unknown samples for which we do not know the dominant hydrometeor type $X_{\text{val}} = \{\boldsymbol{x}_v\}_{v=1}^{N_{\text{val}}}$.

The SVM finds the best linear separation, of type $f(\boldsymbol{x}) = \langle \boldsymbol{w}, \boldsymbol{x} \rangle + b$, for which all training samples are at least at a distance of 1 from the separating plane. In other words, for all training samples, $f(\boldsymbol{x})$ must be greater or equal to 1. To differentiate between positive and negative examples, we also multiply this expression by 1 if the sample is of the positive class and by -1 if it is of the negative class (the two types of hydrometeors). To summarize, the constraint is $y_i(\langle \boldsymbol{w}, \boldsymbol{x}_i \rangle + b) \geq 1, \forall i \in N_{\text{train}}$. The strategy pursued by the SVM (more details in Boser et al., 1992) is to find the separation which maximizes the distance between the closest points of each class, which are also called support vectors. This distance is called the margin and is inversely proportional to the norm of the parameters vector, i.e., $||\boldsymbol{w}||^2$. In order to allow for some classification errors, we also introduce a term $\mathbf{x}i_i$, which is non-zero for samples classified wrongly. The margin maximization problem is the following one:

$$\min_{\boldsymbol{w}, b, \mathbf{x}i} \left\{ \underbrace{\frac{1}{2}||\boldsymbol{w}||^2}_{\text{Complexity of the function}} + \underbrace{C \sum_{i=1}^{N_{\text{train}}} \mathbf{x}i_i}_{\text{Training errors}} \right\} \quad (8)$$

with $\begin{cases} y_i [\langle \boldsymbol{x}, \boldsymbol{w} \rangle + b] \geq 1 - \mathbf{x}i_i \\ \mathbf{x}i_i \geq 0 \quad \text{and} \quad i = 1, \ldots, N_{\text{train}}. \end{cases}$

C is a parameter that controls the constraint of perfect classification: if we allow some errors (by keeping C low), the margin becomes larger, thus reducing the dependence of the final model on training samples, which may be noisy or issued from errors in the measurements. A too high value of C drastically increases the value of the cost function as soon as errors are made. In this case, the resulting model will achieve perfect classification of the training samples, but the risk of over-fitting the training data and achieving poor generalization in the validation phase is higher.

This optimization model is solved using Lagrangian multipliers $\boldsymbol{\alpha}$, which allow us to rewrite the problem as

$$\max_{\boldsymbol{\alpha}} \left\{ \sum_{i=1}^{N_{\text{train}}} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{N_{\text{train}}} \alpha_i \alpha_j y_i y_j \langle \boldsymbol{x}_i, \boldsymbol{x}_j \rangle \right\} \quad (9)$$

s.t. $0 \leq \alpha_i \leq C$ and $\sum_{i=1}^{N_{\text{train}}} \alpha_i y_i = 0$.

When the optimal solution of Eq. (9) is found (i.e., the vector of coefficients $\boldsymbol{\alpha}$), the label of an unknown sample $\boldsymbol{x}_v$ is assigned on the basis of the sign of the decision function, i.e., its position with respect to the hyperplane ($f(\boldsymbol{x}) = 0$):

$$y_v = \text{sign} \left(\sum_{i=1}^{N_{\text{train}}} \alpha_i y_i \langle \boldsymbol{x}_i, \boldsymbol{x}_v \rangle + b \right). \quad (10)$$

In the present formulation, it can be observed that the SVM is only a binary classifier. A number of strategies exist

to reduce multiclass problems to binary problems, and in the present work the one-against-one rule was employed (Hastie and Tibshirani, 1998). The one-against-one rule builds as many binary classifiers as there are pairs of classes. Each classifier is therefore used to assign the time step to one of two possible classes. The time step is eventually classified into the class that received the most assignments.

3.2.2 Nonlinear SVM

The SVM, as has been discussed above, can only solve linear problems (it defines a linear hyperplane). However, there is an elegant solution to solve nonlinear problems. Let us go back to Eqs. (9) and (10): the solution of the optimization does not depend on the training samples themselves but instead only on the dot products between samples (see $\langle \boldsymbol{x}_i, \boldsymbol{x}_j \rangle$ in Eq. 9). In the same way, the prediction for a new sample only depends on its dot products with the training samples (see $\langle \boldsymbol{x}_i, \boldsymbol{x}_v \rangle$ in Eq. 10). Dot products are measures of similarity between the samples. To perform nonlinear classification, we need to find an estimate of their similarity in a projected space of higher dimension $\mathcal{H}$, where linear separation becomes possible[3]. To avoid explicitly defining the coordinates of the samples in the projected space, i.e., $\phi(\boldsymbol{x}_i)$, we can use functions that, even if expressed with points in the original space, correspond to dot products in the projected space $\mathcal{H}$: these functions are called kernels. Without entering mathematical details, which the interested reader can find in Scholkopf and Smola (2001), a kernel corresponds to a similarity function such that $K(\boldsymbol{x}_i, \boldsymbol{x}_j) = \langle \phi(\boldsymbol{x}_i), \phi(\boldsymbol{x}_j) \rangle$. This means that, for a given projection $\phi(\cdot)$, the kernel computed from $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ will correspond to their similarity in the space $\mathcal{H}$ defined by $\phi(\cdot)$. A classification which is linear in the projected space is nonlinear in the original space, as illustrated in Fig. 5.

In practice, in order to obtain a nonlinear classification with the SVM, we replace the dot products in Eqs. (9) and (10) by kernel functions $K(\boldsymbol{x}_i, \boldsymbol{x}_j)$ and $K(\boldsymbol{x}_i, \boldsymbol{x}_v)$, respectively. A classical kernel to obtain such a behavior is the radial basis function (RBF), which is computed as follows:

$$K(\boldsymbol{x}_i, \boldsymbol{x}_j) = \exp \left(\frac{||\boldsymbol{x}_i - \boldsymbol{x}_j||^2}{2\sigma^2} \right). \quad (11)$$

The RBF kernel acts as a Gaussian similarity, which is maximal when considering the same samples ($K(\boldsymbol{x}_i, \boldsymbol{x}_i) = 1$), and decreases jointly with the decrease of similarity between the samples. The bandwidth σ controls the steepness of the Gaussian bell.

3.2.3 SVM-MKL

Even if very successful, SVM remains a black-box model, in the sense that no information about the importance of the

---

[3]The Cover theorem states that the probability of linear separability increases with the dimensionality of the space (Cover, 1965).
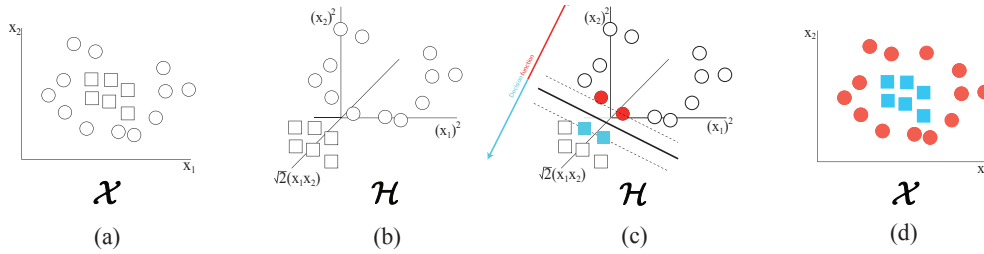
Figure 5. Illustration of the nonlinear SVM. **(a)** A nonlinearly separable data set in the input space $\mathcal{X}$, involving two classes (squares and circles). **(b)** Projection on a 3-D space $\mathcal{H}$ by the kernel $K(\mathbf{x}_i, \mathbf{x}_j) = \langle \mathbf{x}_i, \mathbf{x}_j \rangle^2$. **(c)** Linear classification in the projected space $\mathcal{H}$ (filled dots are support vectors). **(d)** Corresponding nonlinear decision function in the original space $\mathcal{X}$. Adapted from Volpi et al. (2013).

initial variables can be retrieved from its results. All operations are optimized in the projected space $\mathcal{H}$: this means that, while it avoids computation of projection of the samples explicitly, it also prevents the assessment of the importance of the different variables involved. Recent research has offered a solution to this problem by introducing the concept of multiple kernel learning (MKL; Rakotomamonjy et al., 2008).

SVM-MKL builds on the so-called Mercer conditions, stating that a weighted sum of any positive definite function (a requirement for all kernel functions) is again definitely positive (Mercer, 1905). This means that we can design a valid kernel by a linear combination of M base kernels $K_m(\mathbf{x}_i, \mathbf{x}_j)$, each one considering single time step features (in this case $M = 115$) or sets of time step features (in this case $M < 115$ and equals the number of groups of descriptors):

$$K(\mathbf{x}_i, \mathbf{x}_j) = \sum_{m=1}^{M} d_m K_m(\mathbf{x}_i, \mathbf{x}_j). \tag{12}$$

d_m is the weight attributed to each kernel K_m and is a measure of the importance of this kernel in the combination, i.e., of the variables composing it. It usually sums up to 1. The SimpleMKL algorithm proposed in Rakotomamonjy et al. (2008) alternatively optimizes the weights and the SVM and at the same time retrieves the relative importance of each group (d_m), and the SVM model associated with the final weighted combination.

In our experiment, we used SimpleMKL to find the best combination of a series of RBF kernels K_m, each one assigned to a set of features referring to the same particle descriptor ($M = 13$; see Table 1). As an example, K_1 takes into account the eight statistical features (Q10, Q25, Q50, Q75, Q90, IQ75–25, IQ90–10, mean) associated with the hydrometeor fall velocity v descriptor, K_2 the eight features associated with the equivolumetric diameter D_e, and so on.

Table 2. Example of a confusion matrix obtained during validation of the SVM classification for a validation set N_{val}^* of 100 observations. Correct classifications are situated on the diagonal, and misclassifications are in the off-diagonal entries.

					True			
		SP	D	C	G	RIM	AG	MS	R
Predicted	SP	**14**	0	1	3	0	0	0	0
	D	0	**9**	0	0	0	3	0	0
	C	0	0	**9**	0	0	0	0	0
	G	1	0	0	**9**	0	0	0	0
	RIM	0	0	0	0	**8**	0	0	0
	AG	0	0	0	0	0	**11**	1	0
	MS	0	0	0	2	0	1	**13**	0
	R	0	0	0	0	0	0	0	**15**

4 Results and discussion

4.1 Performance assessment metrics

The evaluation of the accuracy of classification is conducted via different metrics. The available N_{train} training observations are divided into two parts (N_{train} and N_{val}^*). N_{train} observations are used as a training set to optimize the SVM parameters C and σ and to train the SVM, while the remaining N_{val}^* observations are kept for validation. A comparison is made between the SVM classification output $\{y_i^*\}_{i=1}^{N_{\text{val}}^*}$, and the true labels $\{y_i\}_{i=1}^{N_{\text{val}}^*}$ by evaluating an 8×8 confusion matrix $\mathbf{C}$, as shown in Table 2. The elements $C(i, j)$ contain the number of observations classified in the ith class, which in reality belong to the jth class. The diagonal contains the correct classifications.

Given the confusion matrix, the global performance of the classifier is quantified by the overall accuracy (OA), and Cohen's kappa ($\mathcal{K}$):

$$\text{OA} = \frac{\sum_{i=1}^{S} C(i, i)}{N} \times 100, \tag{13}$$

$$\mathcal{K} = \frac{\text{OA} - P_{\text{est}}}{1 - P_{\text{est}}} \tag{14}$$

$$P_{\text{est}} = \frac{\sum\limits_{i=1}^{S} \left(\sum\limits_{j=1}^{S} C_{j,i} \sum\limits_{j=1}^{S} C_{i,j} \right)}{N^2},$$

where S is the total number of classes, and N the total number of observations (in our case $S = 8$ and $N = N_{\text{val}}^*$). $\mathcal{K}$ takes into account the correct prediction that might occur by chance, namely P_{est}, and is a robust metric in the case of unbalanced classes.

Then, we look at the performances obtained within each class. For this purpose, we use

$$OA_k = \frac{C(k,k)}{\sum\limits_{i=1}^{S} C(k,i)} \times 100, \tag{15}$$

$$POD_k = \frac{C(k,k)}{\sum\limits_{i=1}^{S} C(i,k)}, \tag{16}$$

$$POFD_k = \frac{\left[\sum\limits_{i=1}^{S} C(k,i) \right] - C(k,k)}{\sum\limits_{i=1}^{S} C(k,i)}, \tag{17}$$

where OA_k is the accuracy of the kth class, and POD_k and $POFD_k$ are respectively the associated probabilities of detection and false detection.

4.2 Evaluation of the quality of the training set

N_{train} observations are available in total as a training set, and we must verify that this amount is sufficient for the present task. In other words, we want to assess here whether a larger N_{train} would significantly improve the hydrometeor classification. To do so we proceeded as follows: (1) $N_{\text{train}} = 400$ was initially randomly split into $N_{\text{train}}^* = 300$ and $N_{\text{val}}^* = 100$; (2) N_{train}^* was iteratively reduced in size, while the original N_{val}^* was kept for validation; (3) evaluation of the performance was conducted at each step; and (4) steps (1)–(3) were repeated with 200 realizations of the original split.

Figure 6 shows the evolution of $\mathcal{K}$ as a function of the number of training samples in the training set (N_{train}^*). We can observe that N_{train}^* larger than 200 did not lead to significant improvements in terms of $\mathcal{K}$, while when N_{train}^* was smaller than 100, the performances started to degrade sharply. These results suggest that the total available labeled samples (400) are sufficient for the present classification task.

4.3 Evaluation of the classification performances

For validation purposes, we now focus on 200 realizations of the case $N_{\text{train}}^* = 300$, $N_{\text{val}}^* = 100$. The classification achieved accurate global results, both in terms of OA and $\mathcal{K}$. As shown in Table 3, $\mathcal{K}$ and OA mean values were 0.88 and 89 %, and in 90 % of the cases they took values higher
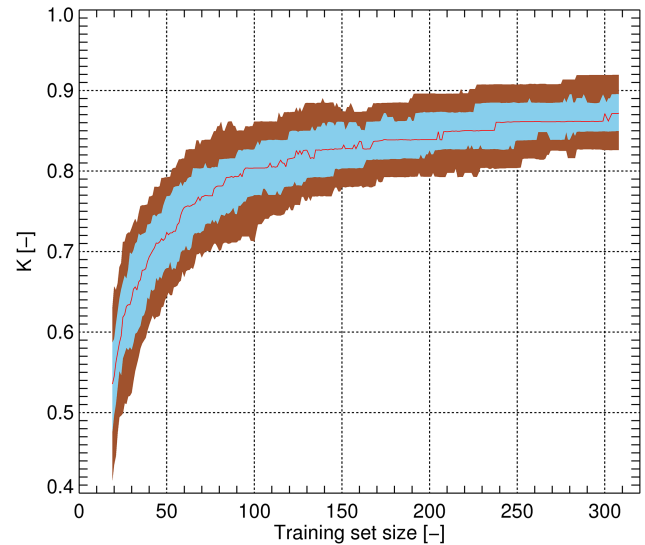

Figure 6. Evolution of $\mathcal{K}$ [–] as a function of the training set size. The solid red line indicates the median, while the blue and brown areas represent Q75–25 and Q90–10, respectively. The size of the training set was varied with a step of 1 between 300 and 20. Statistics are based on 200 realizations.

Table 3. Mean values and relevant quantiles of $\mathcal{K}$ [–] and OA [%] calculated over 200 iterations of the SVM validation procedure.

Parameter	Q10	Q25	Q50	Q75	Q90	Mean
$\mathcal{K}$	0.84	0.86	0.88	0.91	0.93	0.88
OA	86	88	90	92	94	89

than 0.84 and 86 %, respectively. Additionally, $\mathcal{K}$ tended to be close to OA, indicating that correct classification occurring by chance is very unlikely.

The classification performance associated with each hydrometeor class is summarized in Fig. 7. It can be observed that all the hydrometeor classes were identified with median OA_k always greater than 84 %, POD_k greater than 0.84, and $POFD_k$ lower than 0.16. Overall, rainfall (R) hydrometeor class achieved the best scores, together with columns (C). R hydrometeors showed a POD_k equal to 1, meaning that errors for this class were uniquely false detection. In contrast, C hydrometeors showed low $POFD_k$ and OA_k very close to 1, and the errors for this class were due mainly to missed detections, with POD_k with median scores around 0.9. Graupel (G) was mostly affected by missed detections, and showed a relatively large interquantile spread for POD_k, around the median value of 0.87. Small particles (SP) had the highest false detection rate, with median $POFD_k$ close to 0.15. Dendritic snow (D) exhibited the largest interquantile spreads, around otherwise satisfactory median values of 87 % (OA_k), 0.87 (POD_k), and 0.13 ($POFD_k$), followed by rimed particles (RIM) that exhibited a similar behavior, achieving higher scores for all the metrics. Aggregates (AG) and melting snow (MS) were both
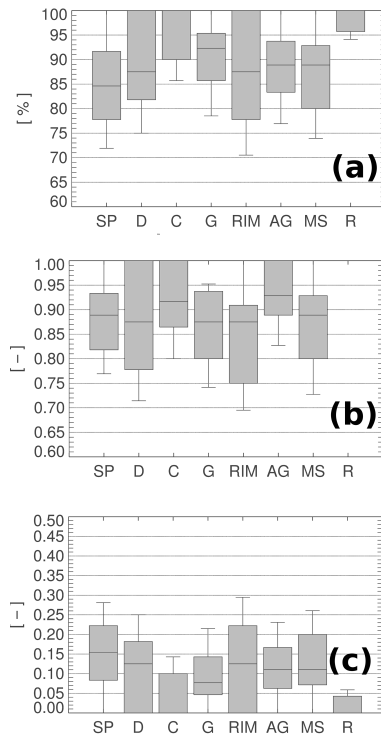
Figure 7. Bar plots of (a) OA_k [%], (b) POD_k [–], (c) $POFD_k$ associated with the eight hydrometeor classes undergoing classification. Statistics were calculated over 200 realizations of the SVM validation.

correctly predicted, with lower interquantile spread, median OA_k larger than 88 %, POD_k larger than 0.88, and $POFD_k$ lower than 0.12.

A last consideration concerns the choice of SVM as a classifier. Other methods are used to solve similar tasks in various fields of the environmental sciences, for example linear discriminant analysis (LDA) or neural networks (NN) (e.g., Goosaert and Alam, 2009; Robert et al., 2013). Comparison with these two methods showed that the proposed SVM scheme outperforms LDA by more than 25% and NN by more than 15% in terms of $\mathcal{K}$.

4.4 Ranking of descriptors

The SimpleMKL algorithm was applied to learn the most relevant descriptors in the classification process, as explained in Sect. 3.2.3. Referring to Eq. (12), it was observed that 5 groups of features out of the 13 (1 per descriptor, each including the 8 or 9 statistical features extracted from its distribution in $\Delta t = 60$ s) accounted for about 70 % (Fig. 8) of the total weights and therefore are considered hereafter as the most important ones. They are, in decreasing order of importance, pixel fraction PF, velocity v, equivolume diameter D_e, form index FORM, and thickness T, with associated weights d_m of 0.193, 0.181, 0.13, 0.112, and 0.098, respectively. This does not imply that the remaining eight
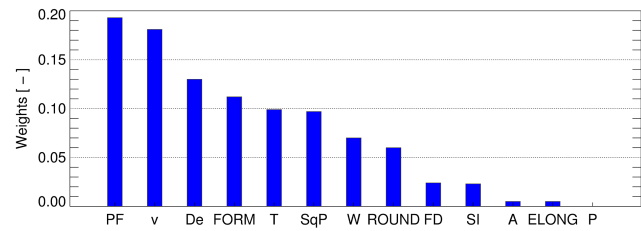

Figure 8. Weights d_m of the 13 K_m kernels, associated with the 13 particle descriptors used in the present study.

descriptors were negligible in the classification process, but that we expect to find a more immediate and intuitive physical meaning in these five top-ranked ones.

5 Application on unlabeled data

This section presents some examples of the classification output, on data not included in the training set of the algorithm, and collected during the measurement campaign in Davos.

5.1 17 March 2011

A snowfall event occurring on 17 March 2011 is presented in Fig. 9. The air temperature recorded in the very close vicinity (less than 50 m away) was constantly below freezing ($\approx -5\,°C$) through the entire event duration, and different ice-phase hydrometeors were identified in the time window shown here. Initially (07:00–09:00 UTC) precipitation was dominated by small particles (SP), followed by a phase of instability (09:00–10:00 UTC) characterized by sharp variations in the identified hydrometeor classes. During the next relatively stable phase (10:00–12:30 UTC), graupel (G) and larger rimed particles (RIM) were identified. Panels (b), (c), and (d) of Fig. 9 illustrate the behavior in time of the three top-ranked particle descriptors, namely pixel fraction PF, equivolume diameter D_e, and fall velocity v. The median PF was around 0.7 during the entire event, indicating relatively high particle compactness. The median D_e was initially below 1 mm (SP phase), and it increased to values between 1 and 2 mm in the latter part of the event characterized mostly by G and RIM classes. v exhibited the same trends as D_e, and it increased when rimed particles and graupel were dominant.

5.2 12 January 2011

A different situation is depicted in Fig. 10, relative to a snowfall event recorded on 12 January 2011. In this case, for the time window shown (19:00–24:00 UTC), precipitation was dominated by aggregates (AG) and dendritic-shaped snow (D, at the end of the event). By comparing the present case with the one shown in Fig. 9, we observe a wider range of variation of particle sizes, with D_e ranging between 0.5 and
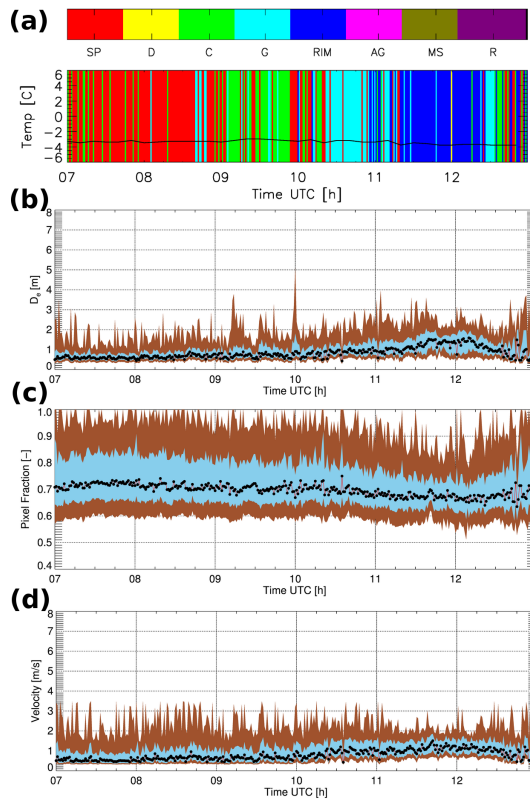
Figure 9. Snowfall event recorded on 17 March 2011. Time series of **(a)** dominant hydrometeor type as classified with the SVM and local ambient temperature [°C], as measured by a closely located (distance ≤ 50 m) weather station; **(b)** D_e [mm]; **(c)** pixel fraction of camera A PF_A; and **(d)** fall velocity v [m s^{-1}]. In panels **(b)**, **(c)**, and **(d)**, black dots connected by the red solid line indicate the median value, while the shaded areas depict Q90–Q10 and Q75–Q25, respectively.

Figure 10. As in Fig. 9 but for 12 January 2011.

8 mm (AG). Particle compactness was lower, with median PF below 0.7 throughout the event, and slightly lower for D than for AG. This is due to the higher geometrical complexity of aggregates and dendrites relative to small particles and graupel. The velocity v did not exhibit peculiar trends, and it remained around values of 1 m s^{-1} in median.

5.3 5 August 2010

The precipitation event that occurred on 5 August 2010 (Fig. 11) illustrates the transition between liquid-phase and ice-phase precipitation well. In the first part of the event (05:00–07:45 UTC) air temperature was around 4 °C, and it dropped to 0 °C in the second part of the event (07:45–08:00 UTC). After 08:00 UTC the air temperature stabilized again around 0 °C. These trends in temperature are directly reflected in the dominant hydrometeor types classified: initially rain (R), then melting snow (MS), and finally aggregates (AG). The rain was characterized by small D_e and $2 \leq v \leq 5$ m s^{-1} (i.e., light rain), which is larger than the
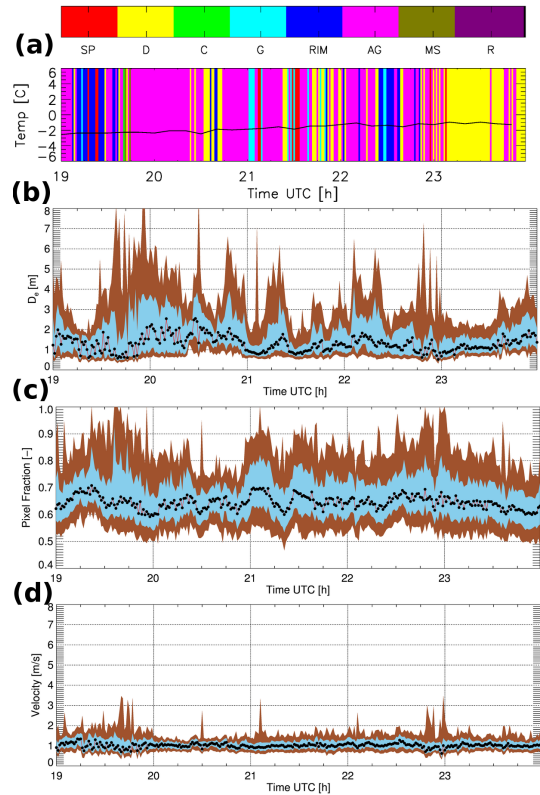
typical velocities of ice-phase hydrometeors, and very high compactness, with a median PF around 0.9. In the transition from R to MS and AG, a clear and relatively smooth trend was observed for the three descriptors shown: v decreased to median values around 1 m s^{-1}; the spread of D_e increased; and the median PF dropped to 0.6 in the AG phase at the same time, as the geometrical complexity of falling hydrometeors increased.

Generally, the transition between R, MS, and AG was well captured in the large available data set. Figure 12 shows the relative number of classifications for each of these three types of hydrometeors as a function of the temperature. Please note that temperature has not been used as an input in the proposed system (Table 1). R always occurred at positive temperatures, MS maximum occurrence was between 2 and 1 °C, and AG around 0 and −1 °C (in agreement with Hobbs et al., 1974). These results give us confidence in the ability of the proposed technique to provide a meaningful classification.

6 Summary and conclusions

In this paper we presented a hydrometeor classification method based on the interpretation of 2DVD data. The classification, conducted with the SVM technique, uses as input the statistical behavior of a set of particle descriptors over
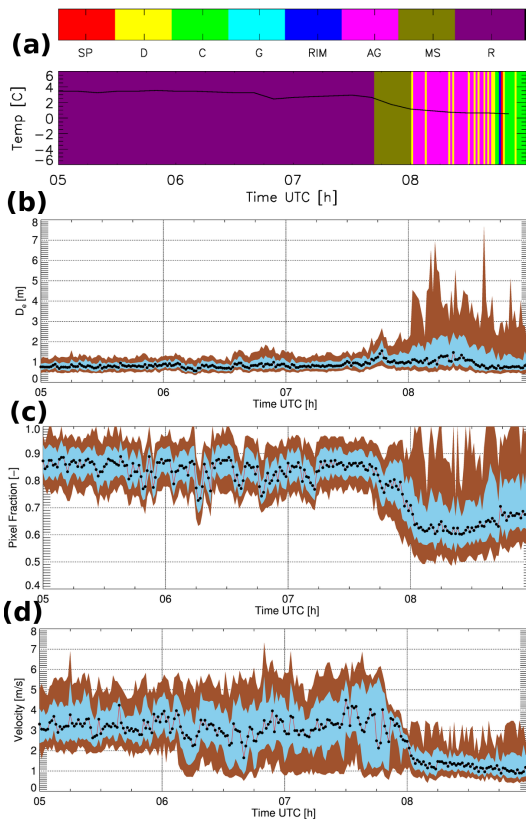
Figure 11. As in Fig. 9 but for 5 August 2010.

Figure 12. Distributions of the occurrence of AG, MS, and R as a function of air temperature. The distributions are obtained by aggregation of all the 2DVD measurements collected during the field experiments in Davos 2009–2011 and Remoray 2012–2013, and temperature data are given by closely located weather stations.

time steps with length $\Delta t = 60$ s. The SVM was trained with 400 examples labeled by expert users, and outputs the dominant hydrometeor type within Δt. Additionally, an estimation of the relative descriptive importance of the input features is provided, which is of particular interest when higher-level information on the particle characteristics is required.

Discrimination is performed between eight hydrometeor classes: small-particle-like, dendrite-like, column-like, graupel-like, rimed-particle-like, aggregate-like, melting-snow-like, and rain. Evaluation of the classification performance was conducted both in global and class-specific terms. The classifier achieved accurate results, with median OA and $\mathcal{K}$ of 90 % and 0.88, respectively. Each of the classes were identified with a median accuracy exceeding 84 %.

Additionally, once trained, the classifier is fast enough to be potentially implemented in real time.

Three classification examples together with the time evolution of the top-ranked particle descriptors were used to illustrate the typical classification products in pure snowfall events and in the transition between snowfall and rainfall. Global hydrometeor-type behavior as well as small-scale fluctuations could be observed.
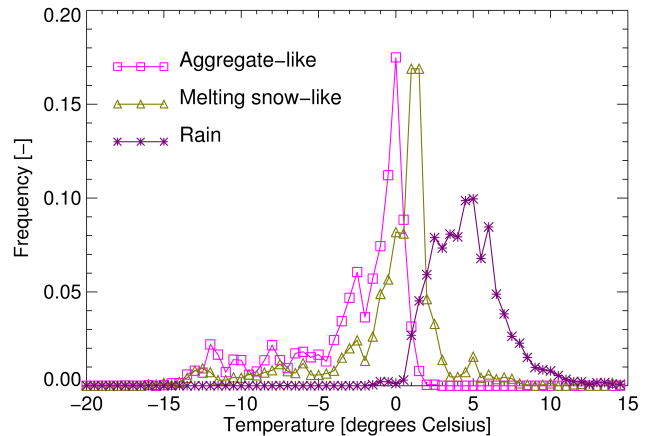
The proposed classification of hydrometeors from the 2DVD measurements provides additional information that can help to better understand the microphysical processes characterizing ice-phase precipitation events. This work has the potential to be a starting point for ground-based quantitative evaluation of products derived from polarimetric weather radars. It can also be adapted and implemented to receive inputs from other particle imaging systems (one or two dimensional), either ground-based or airborne, provided that human interpretation can be carried out for the particle in the training set and that geometrical descriptors can be computed from the particle images.

The main limitation is that the current implementation provides bulk information over a given time step of length Δt, which is large enough to be statistically significant, but cannot provide estimation of hydrometeor mixtures over Δt. Future work will focus on the development of a particle-by-particle classification – which is more challenging in terms of computational requirements – that can lead to explicit quantification of hydrometeor mixtures.

Appendix A: Minimum number of particles for a reliable classification

The proposed classification method employs a set of statistical features calculated over N particles as input, within a time step of length Δt. Thus, when N is small, sampling problems can affect the estimation of such statistics. We want to set a minimum N_{min}, such that if $N > N_{min}$, the classification output is reliable.

Figure A1a shows the contribution that time steps of length Δt with small N had with respect to the total amount of data available, both in terms of total number of particles and in terms of total number of time steps.

We can observe that time steps with low N contribute negligibly to the total particle count, but significantly to the total count of available time steps. In other words, time steps with a low number of particles carry only a small part of the total precipitation, but they are observed frequently.

Figure A1b illustrates the classification performance achieved when $N < 60$. This was obtained by taking random subsets of the available training set (with known labels) and using them as validation of the SVM algorithm trained previously. We can observe that the performance degraded sharply for $N < 20$, and becomes more than 20 % lower than in cases with $N > 60$. A threshold $N_{min} = 20$ was therefore selected.
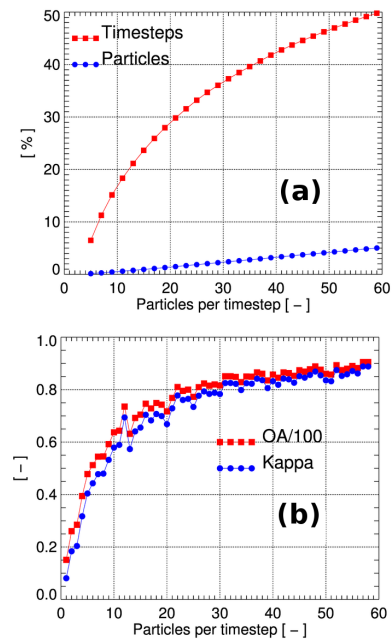

Figure A1. (a) Contributions [%] of time steps of length Δt with less than 60 particles to the total database of observations. **(b)** Classification performance as a function of the number of particles recorded per time step. Time steps with less than three particles did not contribute to these statistics.

References

Beard, K. V.: Terminal velocity and shape of cloud and precipitation drops aloft, J. Atmos. Sci., 33, 851–864, 1976.

Boser, B., Guyon, I., and Vapnik, V.: A training algorithm for optimal margin classifiers, in: 5th ACM Workshop on Computational Learning Theory, 144–152, Pittsburgh, USA, 1992.

Brandes, E. A., Ikeda, K., Zhang, G., Schonhuber, M., and Rasmussen, R. M.: A statistical and physical description of hydrometeor distributions in Colorado snowstorms using a video disdrometer, J. Appl. Meteor. Clim., 46, 634–650, doi:10.1175/JAM2489.1, 2007.

Camps-Valls, G. and Bruzzone, L.: Kernel-based methods for hyperspectral image classification, IEEE T. Geosci. Remote Sens., 43, 1351–1362, doi:10.1109/TGRS.2005.846154, 2005.

Camps-Valls, G. and Bruzzone, L. (Eds.): Kernel Methods for Remote Sensing Data Analysis, Wiley, 2009.

Cao, Q., Zhang, G. F., Brandes, E., Schuur, T., Ryzhkov, A., and Ikeda, K.: Analysis of video disdrometer and polarimetric radar data to characterize rain microphysics in Oklahoma, J. Appl. Meteor. Clim., 47, 2238–2255, doi:10.1175/2008JAMC1732.1, 2008.

Chandrasekar, V., Keranen, R., Lim, S., and Moisseev, D.: Recent advances in classification of observations from dual polarization weather radars, Atmos. Res., 119, 97–111, doi:10.1016/j.atmosres.2011.08.014, 2013.

Cover, T. M.: Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition, IEEE Transactions on Electronics and Computers, EC-14, 326–334, 1965.

Delanoe, J., Protat, A., Jourdan, O., Pelon, J., Papazzoni, M., Dupuy, R., Gayet, J. F., and Jouan, C.: Comparison of airborne in situ, airborne radar-lidar, and spaceborne radar-lidar retrievals of polar ice cloud properties sampled during the POLARCAT campaign, J. Atmos. Oceanic Technol., 30, 57–73, doi:10.1175/JTECH-D-11-00200.1, 2013.

Dolan, B. and Rutledge, S. A.: A theory-based hydrometeor identification algorithm for X-band polarimetric radars, J. Atmos. Oceanic Technol., 26, 2071–2088, doi:10.1175/2009JTECHA1208.1, 2009.

Elmore, K. L.: The NSSL hydrometeor classification algorithm in winter surface precipitation: evaluation and future development, Weather Forecast., 26, 756–765, doi:10.1175/WAF-D-10-05011.1, 2010.

Feind, R. E.: Comparison of three classification methodologies for 2D probe hydrometeor images obtained from the armored T-28 aircraft, Tech. Rep. SDSMT/IAS/R08-01, Institute of Atmo-

spheric Sciences, South Dakota School of Mines and Technology, Rapid City, SD, USA, 2008.

Foresti, L., Tuia, D., Kanevski, M., and Pozdnoukhov, A.: Learning wind fields with multiple kernels, Stoch. Env. Res. Risk A., 25, 51–66, doi:10.1007/s00477-010-0405-0, 2011.

Goosaert, E. and Alam, R.: Ensemble classifier for winter storm precipitation in polarimetric radar data, in: Seventh conference on artificial intelligence and its applications to the environmental sciences, Phoenix, USA, 2009.

Hanesch, M.: Fall velocity and shape of snowflakes, Ph.D. thesis, Swiss Federal Institute of Technology Zurich, 1999.

Hastie, T. and Tibshirani, R.: Classification by pairwise coupling, Ann. Stat., 26, 451–471, 1998.

Hobbs, P. V., Chang, S., and Locatelli, J. D.: The dimensions and aggregation of ice crystals in natural clouds, J. Geophys. Res., 79, 2199–2206, doi:10.1029/JC079i015p02199, 1974.

Houze, R. A. J.: Cloud dynamics, Academic Press, San Diego, 1993.

Huang, G., Bringi, V. N., Cifelli, R., Hudak, D., and Petersen, W. A.: A methodology to derive radar reflectivity liquid equivalent snow rate relations using C-band radar and a 2D video disdrometer, J. Atmos. Oceanic Technol., 27, 637–651, 2010.

Jiao, L. and Liu, Y.: Analyzing the shape characteristics of land use classes in remote sensing imagery, ISPRS Annals of Photogrammetry, Remote Sens. Spatial Inf. Sci., I-7, 135–140, doi:10.5194/isprsannals-I-7-135-2012, 2012.

Kruger, A. and Krajewski, W. F.: Two-dimensional video disdrometer: a description, J. Atmos. Oceanic Technol., 19, 602–617, 2002.

List, R. and Schemena, R.: Free-fall behavior of planar snow crystals, conical graupel and small hail, J. Atmos. Sci., 28, 110–115, doi:10.1175/1520-0469(1971)028<0110:FFBOPS>2.0.CO;2, 1971.

Löffler-Mang, M. and Joss, J.: An optical disdrometer for measuring size and velocity of hydrometeors, J. Atmos. Oceanic Technol., 17, 130–139, 2000.

Marzano, F. S., Cimini, D., and Montopoli, M.: Investigating precipitation microphysics using ground-based microwave remote sensors and disdrometer data, Atmos. Res., 97, 583–600, doi:10.1016/j.atmosres.2010.03.019, 2010.

Mercer, J.: Functions of positive and negative type and their connection with the theory of integral equations, Philos. Tr. R. Soc., CCIX, 215–228, 1905.

Nespor, V., Krajewski, W. F., and Kruger, A.: Wind-induced error of raindrop distribution measurement using a two-dimensional video disdrometer, J. Atmos. Oceanic Technol., 17, 1483–1492, 2000.

Rakotomamonjy, A., Bach, F., Canu, S., and Grandvalet, Y.: SimpleMKL, J. Mach. Learn. Res., 9, 2491–2521, 2008.

Robert, S., Foresti, L., and Kanevski, M.: Spatial prediction of monthly wind speeds in complex terrain with adaptive general regression neural networks, Int. J. Climatol., 33, 1793–1804, doi:10.1002/joc.3550, 2013.

Schneebeli, M., Dawes, N., Lehning, M., and Berne, A.: High-resolution vertical profiles of polarimetric X-band weather radar observables during snowfall in the Swiss Alps, J. Appl. Meteor. Clim., 52, 378–394, doi:10.1175/JAMC-D-12-015.1, 2013.

Scholkopf, B. and Smola, A. J.: Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond, MIT Press, Cambridge, MA, USA, 2001.

Schönhuber, M., Lammer, G., and Randeu, W. L.: The 2D video disdrometer, Chapter 1 in: Precipitation: Advances in Measurement, Estimation and Prediction, edited by: Michaelides, S., 2008.

Schuur, T. J., Ryzhkov, A. V., Zrnic, D. S., and Schonhuber, M.: Drop size distributions measured by a 2D video disdrometer: Comparison with dual-polarization radar data, J. Appl. Meteor., 40, 1019–1034, doi:10.1175/1520-0450(2001)040<1019:DSDMBA>2.0.CO;2, 2001.

Shupe, M. D.: A ground-based multisensor cloud phase classifier, Geophys. Res. Lett., 34, L22809, doi:10.1029/2007GL031008, 2007.

Straka, J. M., Zrnic, D. S., and Ryzhkov, A. V.: Bulk hydrometeor classification and quantification using polarimetric radar data: synthesis of relations, J. Appl. Meteor., 39, 1341–1372, 2000.

Sullivan, S.: Evaluation of support vector machines and minimax probability machines for weather prediction, in: Seventh conference on artificial intelligence and its applications to the environmental sciences, Phoenix, USA, 2009.

Thurai, M. and Bringi, V. N.: Drop axis ratios from a 2D video disdrometer, J. Atmos. Oceanic Technol., 22, 966–978, doi:10.1175/JTECH1767.1, 2005.

Thurai, M., Hudak, D., and Bringi, V. N.: On the possible use of copolar correlation coefficient for improving the drop size distribution estimates at C band, J. Atmos. Oceanic Technol., 25, 1873–1880, doi:10.1175/2008JTECHA1077.1, 2008.

Thurai, M., Szakall, M., Bringi, V. N., Beard, K. V., Mitra, S. K., and Borrmann, S.: Drop shapes and axis ratio distributions: comparison between 2D video disdrometer and wind-tunnel measurements, J. Atmos. Oceanic Technol., 26, 1427–1432, doi:10.1175/2009JTECHA1244.1, 2009.

Tuia, D., Camps-Valls, G., Matasci, G., and Kanevski, M.: Learning relevant image features with multiple-kernel classification, IEEE T. Geosci. Remote Sens., 48, 3780–3791, doi:10.1109/TGRS.2010.2049496, 2010.

Volpi, M., Tuia, D., Bovolo, F., Kanevski, M., and Bruzzone, L.: Supervised change detection in VHR images using contextual information and support vector machines, Int. J. Appl. Earth Obs. Geoinf., 20, 77–85, doi:10.1016/j.jag.2011.10.013, 2013.

Xue, M., Droegemeier, K. K., and Wong, V.: The Advanced Regional Prediction System (ARPS); A multi-scale nonhydrostatic atmospheric simulation and prediction model. Part I: model dynamics and verification, Meteorol. Atmos. Phys., 75, 161–193, doi:10.1007/s007030070003, 2000.

Yuter, S. E., Kingsmill, D. E., Nance, L. B., and Loeffler-Mang, M.: Observations of precipitation size and fall speed characteristics within coexisting rain and wet snow, J. Appl. Meteor. Clim., 45, 1450–1464, doi:10.1175/JAM2406.1, 2006.

Zeng, J. and Qiao, W.: Support vector machine-based short-term wind power forecasting, in: Power Systems Conference and Exposition (PSCE), 2011 IEEE/PES, 1–8, doi:10.1109/PSCE.2011.5772573, 2011.

Zhang, G., Xue, M., Cao, Q., and Dawson, D.: Diagnosing the intercept parameter for exponential raindrop size distribution based on video disdrometer observations: model development, J. Appl. Meteor. Clim., 47, 2983–2992, doi:10.1175/2008JAMC1876.1, 2008.

Zhang, G., Luchs, S., Ryzhkov, A., Xue, M., Ryzhkova, L., and Cao, Q.: Winter precipitation microphysics characterized by polarimetric radar and video disdrometer observations in central oklahoma, J. Appl. Meteor. Clim., 50, 1558–1570, doi:10.1175/2011JAMC2343.1, 2011.