

Convolutional neural networks for specific and merged data sets of optical array probe images: compatibility of retrieved morphology-dependent size distributions

Louis Jaffeux, Jan Breiner, Pierre Coutris, and Alfons Schwarzenböck

Laboratoire de Météorologie Physique (UMR6016)/UCA/CNRS, Aubière, France

Correspondence: Louis Jaffeux (louis.jaffeux@uca.fr)

Received: 21 June 2024 – Discussion started: 18 September 2024 Revised: 24 February 2025 – Accepted: 4 March 2025 – Published: 3 June 2025

Abstract. This study addresses the challenges of ice particle morphology classification from images of optical array probes (OAPs) and proposes a more refined processing to enable better interpretation of observational data. The convolutional neural network methodology is applied to train classification tools for hydrometeor images from optical array probes. Two models were developed in a previous work for the Precipitation Imaging Probe (PIP) and 2D Stereo (2D-S). In addition, three new models are introduced in this study: one for the Cloud Imaging Probe (CIP), one for the High Volume Particle Spectrometer (HVPS), and a global model trained on a data set that merges all available data from the above four instruments. The methodology of retrieving morphology-specific size distributions from OAP data is provided. Size distributions for each morphological class, obtained using both the specific and global classification models, are compared for the data set of the ICE-GENESIS (Creating the next generation of 3D simulation means for icing) project, where all four probes were operated simultaneously. The reliability and coherence of these newly obtained machine learning classification tools are clearly demonstrated. The analysis shows significant advantages of using the global model compared to the specific ones. The presented methodology retrieves morphology-specific crystal size distributions that effectively allow systematic identification of microphysical growth processes from OAP data sets mainly collected on research aircraft during measurement campaigns. By combining the quantitative reliability of OAP-derived total number and mass size distributions with advanced machine learning morphological individual crystal classification, this approach establishes a foundation for investigations on past OAP data sets to be reinterpreted.

1 Introduction

The topic of ice crystal shapes has stimulated the imagination of cloud enthusiasts for centuries. The morphology of solid hydrometeors is closely related to their growth history. These shapes reflect the occurrence of certain microphysical processes inside ice clouds and, in some cases, even pinpoint their location in time and space (Pasquier et al., 2023). Particle morphology, therefore, helps explaining different pathways to form atmospheric ice. In particular, the shape of solid hydrometeors may be used to deduce environmental conditions, such as temperature, humidity, and turbulence levels, that influence crystal growth within clouds.

The 3D particle geometry accounts for numerous properties of ice particles, such as fall velocity (Vázquez-Martín et al., 2021; Locatelli and Hobbs, 1974), capacitance (Westbrook et al., 2008), scattering properties (Wyser, 1999), or melting behavior (Matsuo and Sasyo, 1981; Knight, 1979). In addition, several phenomena arise from the natural differentiation of hydrometeors. Extreme precipitation rates, electrification, and extended cloud lifetimes are some of the visible consequences of ice particle interactions. These cloudscale features ultimately play important roles in the global climate system and feedback on atmospheric conditions and thus cloud formation itself. The intricate nature of these feedback mechanisms and their role in the emergence of new cloud-scale properties are illustrated in the next paragraph to emphasize the importance of morphology.

Among hydrometeor types, dendritic crystals stand out for their high capacitance, a property arising from their specific 3D shape. They grow in the temperature region $(-15 \text{ }^{\circ}\text{C})$ where supersaturation reaches its potential maximum (Pruppacher and Klett, 2010). Moreover, their fall velocity is relatively low (Fukuta and Takahashi, 1999), resulting in extensive residence times in their original growth region. Thus, dendritic crystals are draining significant amounts of water vapor. In contrast, graupel particles fall rapidly and have rimed surfaces that inhibit further depositional growth (Jensen and Harrington, 2015). In addition, they are particularly resistant to melting. As a result, rimed particles sediment more effectively towards the ground, draining the cloud of some of its condensed water. When graupel and dendrites collide with one another, charge transfer happens (Emersic and Saunders, 2010). The charging events from these collisions lead to cloud electrification and can eventually trigger lightning. Lastly, lightning strikes play important roles in the climate system, such as the ignition of wildfires that release massive amounts of gas species (including greenhouse gases) into the atmosphere (Knorr et al., 2017). This example shows that different hydrometeor types contribute more effectively to different cloud processes, such as dissipation of supersaturation for dendrites and precipitation for graupel particles. In the presented example, these processes partly control the cloud life cycle and fallout rate, two key characteristics of clouds. In addition, the existence of differentiated particle types allows for the explanation of further phenomena, such as lightning. Because ice morphology is central to ice cloud mechanisms, habit classification of ice particles is, therefore, essential to improving weather and climate models.

For airborne in situ measurements, morphology is often limited to qualitative information because of instrumental limitations (Zhu et al., 2015; McFarquhar et al., 2007). However, quantitative size distributions are produced by analyzing data obtained with optical array probes (OAPs). Historically, the low resolution of OAP images made it difficult to extract morphological information using feature-based approaches. Manual classification was limited due to the large volume of images produced by these probes. However, the eye of an experienced microphysicist is able to distinguish particle shapes and assign the proper morphology to most of the produced 2-D images. Feature-based approaches have tried to capture this human skill using relevant geometric characteristics of 2-D hydrometeor images (Duroure, 1982; Rahman et al., 1981) with limited success and slow processing speed (Praz et al., 2017).

In atmospheric science, the integration of machine learning (ML) techniques has already demonstrated transformative potential, from predicting weather patterns to analyzing cloud organizations in satellite imagery. When it comes to classifying ice particles, ML offers significant advantages over traditional methods, such as manual classification or feature-based algorithms. These conventional approaches often struggle with reliability, the sheer volume of data, and subjective biases, whereas ML provides scalability, consistency, and objectivity. With the advent of artificial intelligence, the human ability to recognize shapes can finally be emulated (Krizhevsky et al., 2012) with a neural network architecture called convolutional neural networks (CNNs). This type of network applies 2-D filters to input images and subsequently created feature maps, thereby creating hierarchical complex abstractions of the image based on the data sets they were trained with. In between these convolutions, pooling layers are used in order to summarize and reduce the size of the obtained feature maps. This step enables strong generalization capabilities, which are required to account for the high variability that is inherent to ice particle shapes and orientations. Finally, fully connected layers are used to combine the final abstractions of the initial image with a category, learned from its manually labeled training data. CNNs are particularly well-suited for image recognition tasks because they automatically learn to identify patterns and features, without requiring human-defined rules, which are the cause for the bad generalization of traditional feature-based approaches.

Using CNNs, Przybylo et al. (2021) and Schmitt et al. (2024) have developed classification tools for other instruments that record ice particle shapes: the Cloud Particle Imager (CPI), which is a high-resolution CCD imager, and the Particle Phase Discriminator mark 2 and the Small Ice Detector, which are 2-D scattering probes. CPI images are highly valuable due to their high resolution and 256 grayscale levels. In contrast, optical array probes (OAPs) produce larger data sets of particle images because of their much larger sample volumes. However, their lower resolution and the fact that those images are mainly black-and-white images also pose significant challenges for traditional classification techniques. Using CNN algorithms, Wu et al. (2020), Jaffeux et al. (2022), and Zhang et al. (2023) achieved reliable automatic classification with single probe data sets. These studies demonstrated that with a sufficiently large and welllabeled data set, CNNs can achieve accuracy comparable to that of microphysics specialists in recognizing the complex shapes of ice particles for OAP images. The current study extends the previously developed CNN models, in order to retrieve a most realistic description of the sampled environments. This is achieved by combining OAP-derived quantitative concentration measurements, which have historically provided reliable data for particle concentrations and size distributions, with advanced automatic classification tools to enhance morphological analysis. These newly obtained morphology-specific size distributions are a major step in understanding the dynamics of the interactions between different ice particle populations that grew under different microphysical regimes and that were clearly identified as the source for related cloud-scale phenomena.

While classification methodologies have been developed, comparisons of coherent observations from different imaging instruments remain scarce. This research gap presents an opportunity to explore data sets and validate classification tools. Multiple OAPs are often mounted on research aircraft in order to cover complementary size ranges with some overlap. This setup ensures instrumental redundancy and enables comparisons, while also capturing the broadest possible spectrum of particle sizes. This study introduces a methodology for splitting particle size distributions into morphology-specific distributions for respective OAP probes, enabling multi-probe comparisons to be performed for each hydrometeor category. These comparisons not only validate classification tools but also reveal the ability of each probe to capture morphological details of ice particles across different sizes and shapes. Furthermore, a new classification tool is trained in this study and used to process images from four different OAP instruments with different pixel resolution and size ranges, while all previous works on OAP CNNs only used single probe data sets. The obtained spectra cover 2 orders of magnitudes in size, consisting in most of the upper range of the particle size spectrum (from 300 µm up to 1.9 cm), and overlap largely in the case of the ICE-GENESIS data set, which contains OAP data sampled simultaneously with the four most used OAP instruments (2D Stereo (2D-S), High Volume Particle Spectrometer (HVPS), Cloud Imaging Probe (CIP), and Precipitation Imaging Probe (PIP)). This instrumental setup, combined with the machine learning classification tools, provides a first basis to analyze and compare habit-dependent particle size distributions over large size ranges. The use of the global CNN model versus CNN models limited to solely probe-specific data sets is a key focus of the article.

The procedure that leads to quantitative estimates of morphology-specific particle size distributions for different OAP instruments is detailed in the first section. First, two already existing classification tools, one for the 2D-S and one for the PIP, are presented, tested, and improved for the ICE-GENESIS data set. Then, three additional CNNs that were developed for the CIP, HVPS, and all four probes simultaneously are quickly presented. Finally, the extracted morphological information and the quantitative bin concentration estimation are combined. In a second part, the data set is briefly presented in terms of the encountered cloud conditions (e.g temperature range, cloud depth, and cloud origins) and total OAP data. After that, the compatibility and coherence of the five classification tools are explored on this data set by studying morphology-specific size distributions for each class. The final section provides a summary of the content of the article; conclusions on the size distribution analysis, including potential improvements and recommendations; and finally, a discussion on the benefits of the use of the developed tools for the scientific community.

2 Methodology

CNNs are able to solve the problem of shape recognition by reaching human levels for single images (Krizhevsky et al., 2012). Nonetheless, these algorithms are not flawless, especially because they rely on manually gathered data sets. Faulty class attributions may stem from the arbitrary definition of image classes or the possibility of encountering undefined particle types in acquired data. Understanding that CNNs are black-box algorithms whose mistakes are difficult to decipher is the main motivation for their testing. It is indeed necessary to address this weakness and evaluate the quality of the extracted morphological data. Jaffeux et al. (2022) developed two classification algorithms for the Precipitation Imaging Probe (PIP) and the 2D-S probes. To these two instruments, the CIP and HVPS have been added.

The CNN structure used in this study is similar to the one of the AlexNet model (Krizhevsky et al., 2012) (as shown in Fig. 1). It consists of two parts: a feature extractor and a classifier. The first part takes a fixed size image as input and follows a hierarchical structure of successive convolution and subsampling layers. This part converts the initial image and the subsequently created feature maps into smaller, summarized, higher level feature maps. The convolution layer applies the dot product to the values of each pixel and its surrounding in a 3 by 3 square and 3 by 3 filters (also called kernels), which are trained to account for abstracted features. Then, the subsampling layer reduces the obtained feature map to its more crucial information using a 2 by 2 max pooling filter, thus dividing the number of pixels of the feature map by 4. Since these operations do not result in an increase in computational cost, the number of filters applied within each convolution layer is accordingly doubled as the feature map reaches deeper levels of the feature extractor. The twolayer types are applied one after the other until the size of the feature maps reaches one. The number of convolution layers depends therefore on the size of the initial image. In the case of the ones obtained with the 2D-S and HVPS, which both have 124 photodiodes, a size of 200 by 200 was used, resulting in six convolution layers. In the case of the CIP and PIP probes, which both have 64 photodiodes, a size of 110 by 110 was used, resulting in five convolution layers.

Finally, the classifier, a fully connected perceptron with a single hidden layer, is used to attribute classes to the combination of the most abstracted features that were extracted. During the training, 20 % of the labeled data are randomly taken out for the final testing, 16 % are used for validation during training, and 64 % are used for training the filter weights and the synapses from the fully connected layers. The images are padded to the adequate input size. Then, they undergo a random flip operation (vertically, horizontally, both, or none) in order to produce more variety in the orientation of the particles without requiring pixel interpolation. Bayesian parameter optimization is performed for hyperparameter tuning, including the number of neurons used in the classifier dropout (Srivastava et al., 2014) for every convolution and the fully connected layers.

The goal of this first section is to describe the training data and the morphological classes associated with each CNN model and to report the results of the training using confusion matrices and training reports obtained on test data sets.



Figure 1. Structure of the convolutional neural network used in this study. For more details, see Jaffeux et al. (2022).

2.1 Morphological classes

When considering clouds and their potential to produce precipitations, ice particles can be separated into three general types: pristine crystals, intermediary particles, and ultimate precipitating particles. Their respective number and mass size distributions and concentrations reflect the strength of the processes governing their appearance, growth and consumption. Because revealing and measuring these effects is the end goal of the developed tools, the classes defined in this subsection fall into these three categories.

- 1. Pristine crystals are single crystals formed through the initiation of ice particles that grow by the deposition of vapor. Pristine crystals disappear through self-aggregation, scavenging by other particle types, and shape alteration via riming or secondary deposition regimes (in the special case of capped columns). These particles include plates and dendrites that are designed under a common class named hexagonal planar crystals (HPCs). Because plates and dendrites grow in adjacent environments (Fukuta and Takahashi, 1999), they are considered a continuum here. Columns and needles (Co) are the other types of pristine crystals defined for the models.
- 2. Intermediary particles are formed from pristine crystals and the intermediary particles themselves. They grow through the aggregation of pristine crystals and intermediary particles, including self aggregation and secondary deposition. They are consumed by collection and riming. Based on the available OAP data sets, combination of bullets and columns (CBCs), complex assemblages (CAs), fragile aggregates (FAs), and capped columns (CCs) are the defined classes corresponding to this intermediary type. Combinations of bullets and combinations of columns are hardly differentiable with the coarse resolution and binary nature of OAP images. Both particle types follow the definition of intermediary types; for this reason, they were put together. Particles with complex shapes and sharp edges, which exhibit

transparency and are often composed of spatial plates, are commonly obtained but only with the 2D-S. The corresponding particles likely formed through the aggregation of pristine crystals and/or significantly grew by deposition in different environments. These observations motivated the definition of the CA class for the 2D-S. In some cases, the individual elements composing an aggregate cannot be identified, either because these elements are too small with respect to the pixel resolution or because the elements are individually amorphous (e.g., aggregates of ice fragments). The FA class corresponds to these aggregate types in cases where the bounds between the monomers are relatively thin.

3. Ultimate precipitating particles are formed and grow by the riming and aggregation of any type of crystal, including self aggregation. The mass fall rate (or downward mass flux) of ultimate particles is thereby roughly constant. Two classes are defined to correspond to this definition: compact particles (CPs) and rimed aggregates (RAs). Both of these morphological classes are close in terms of shapes and mostly designate two archetypes of dense ice particles, CPs being the most compact and RAs being defined as particles that visibly contain several monomers.

All the classes defined above are summarized in Fig. 2 with actual examples from the training data for each of the four probes. It can be noted that for the 2D-S and CIP, a water droplet (WD) class was added. In addition to these classes, two artifact classes were defined for each of the SPEC instruments: a diffracted class for the 2D-S (Dif) and a fragmented particle class for the HVPS (FP). These classes confirm the fact that despite the use of inter-arrival time algorithm (Field et al., 2006), and treatment of diffracted images (Vaillant de Guélis et al., 2019), some of these artifact images are still found within OAP images. All training data sets and codes for the training can be found in the GitHub repository (see https://doi.org/10.5281/zenodo.15573218, Jaffeux, 2025). The 2D-S, CIP, PIP, and HVPS training data sets are composed of 6561, 5163, 3281, and 4290 images,

Class names	2DS examples (width = 2 mm)			CIP examples (width = 2.75 mm)			PIP examples (width = 11 mm)			HVPS examples (width = 19.2 mm)		
Compact Particles							*		•	•	*	6 :
Fragile Aggregates	K *	**	\$	¥	*****	яř	k g r	¥	*** ***	*	a.¢*	**
Columns and Needles	1	1		-		,	-	١	١	`	1	\setminus
Hexagonal Planar Crystals	Ø	25	*	*		*	*	×	÷	*	*	*
Rimed Aggregates										*	Å	*
Rimed Aggregates Combinations of Bullets and Columns	*	¥	¥	K		*		*	**	*	ri de la constancia de	\$
Rimed Aggregates Combinations of Bullets and Columns Complex Assemblages	+	X	\¥ ₽	K	×	*	*	*	**	*	raft R	\$
Rimed Aggregates Combinations of Bullets and Columns Complex Assemblages Capped Columns	↓ → 3	r ×	 ✓ ✓ 	k	×	*		*	**	*	rî.	*

Figure 2. The nine morphological classes defined for the classification algorithms for the 2D-S, CIP, and PIP (Jaffeux et al., 2022).

respectively. The building of these training data sets is the result of an iterative process of training, testing, and gathering more pertinent data. An appreciable number of images in each class is a not a necessary condition to allow the CNN model to train and, more importantly, generalize successfully.

As mentioned in the Introduction, the entirety of the training data for each probe was gathered into a single data set with the aim to train a model adapted for the data of any OAP. Two noticeable adjustments were made in the making of this new data set: additional 2D-S dendrite and hollow column images were added, capped columns for the PIP were added, and the RA images were included in the CP class. The CA, Dif, and FP morphological classes only include data from one probe (2D-S for the CA and Dif and HVPS for SP). The consequences of the decision to keep these classes in a general model are discussed in the analysis section. In total, 21390 images constitute this "global" data set. This number can be put into perspective with the 24 720, 9000, and 33 300 images used in Przybylo et al. (2022), Schmitt et al. (2024), and Zhang et al. (2023), respectively, and which used transfer learning for similar number of classes with the same objective of classifying ice particle shapes.

Hereafter, the quality of the training of each CNN, associated with each data set, is demonstrated. The CIP, HVPS, and "global" data sets are essentially built with data from the ICE-GENESIS campaign. For the three corresponding CNNs, the confusion matrices and training reports are shown and described briefly. For the 2D-S and PIP algorithms, which were obtained in Jaffeux et al. (2022) with data acquired prior to the ICE-GENESIS campaign, additional specific testing and assimilation are presented instead.

2.2 Additional CNNs: the Cloud Imaging Probe (CIP), High Volume Particle Spectrometer (HVPS), and global model

2.2.1 CIP

The CIP has a 25 µm resolution and 64 pixels, providing measurements of hydrometeors in the 12.5 to 1600 µm range. As for any OAP instrument, the upper size limit can be extended at the cost of measurement statistics and bias in undersizing truncated particles, when reconstructing those. The 2D-S classes are used except for the CA class, which was characterized by some level of transparency which could not be found within available CIP images. The same training methodology was used for the CIP CNN as for the PIP and 2D-S ones; see Jaffeux et al. (2022). The CIP CNN was trained exclusively on ICE-GENESIS data, as opposed to the PIP and 2D-S CNNs. The associated training confusion matrix and training report can be found in Fig. 3. The most noticeable confusion happens between CPs and HPCs, but this confusion happens at the expense of the precision of the CP class, which is a good outcome considering the reasoning given in the previous subsection. The total accuracy is slightly below 90%. It can be noted that both Co and WD classes are very well recognized.

2.2.2 HVPS

The HVPS has 150 μ m resolution and 128 pixels, providing measurements of hydrometeors in the 75 to 19 200 μ m range. For the HVPS, all defined PIP classes are used, with the addition of the FP class. The HVPS CNN was trained exclusively on ICE-GENESIS data, similarly to the CIP CNN. The resulting confusion matrix and training reports are shown in Fig. 4. Few HPCs were found within the HVPS data sets (318), leading to low precision for this class. The most noticeable confusion happens between aggregate classes, in particular the CNN identifies some RAs and CBCs as FAs. The total accuracy is around 85 %.

2.2.3 Global model

On the validation set, accuracy reached 96.8 %, and testing resulted in the confusion matrix and training reports shown in Fig. 5. Relatively low confusion was found between the FA, CBC, and CP classes on the one hand and between CPs and HPCs on the other hand. The quality of the obtained model matches that of the specific models. Consistently with the specific models of 2D-S and CIP, Co and WD classes are well recognized in the test data set.

2.3 CNN-specific test and data assimilation

In order to specifically evaluate the performance of the previously trained algorithms (for PIP and 2D-S) and improve their representation for each class, 200 images were manually extracted from the whole ICE-GENESIS data set for each class for the PIP and 2D-S, in order to test both CNNs with respect to their accuracy concerning each morphological class. If the algorithm shows significant amounts of error, data assimilation has to be performed to hopefully resolve the issues. The authors strongly encourage manual inspection by the reader of the 200 images extracted for each class as their selection is inherently subjective. Such manual verification is required to help in understanding the predictions of the CNN models in subsequent analysis.

2.3.1 2D-S

Figure 6a describes the results of the specific testing. While columns are perfectly recognized, and generally most classes are well identified, some porosity between CBCs and FAs, on the one hand, and between HPCs and CPs, on the other hand, is found in reasonably limited amounts. However, the model iteration completely fails to recognize WD. Comparison between the test data and training data shows that the test droplets are vertically elongated, while the training ones were either horizontally elongated or spherical. This difference stems from the processing methodology used at the

Laboratoire de Météorologie Physique (LaMP), which is the setup of a high default true air speed (TAS) for OAPs rather than an online direct update from the plane's central computer. While size parameters are corrected, since the default TAS is higher than the real one, images are not resized accordingly. In other words, depending on the plane's speed, different levels of 2-D image deformation were experienced for the droplets. Since water droplets have very characteristic shapes, the CNN is highly disturbed by this change and does not recognize water droplets from the ICE-GENESIS data at all. Finding a remedy to this problem is required, not only because identifying WD is essential to the microphysical analysis but also because their presence currently harms the precision of the CP and HPC classes.

The tested data were assimilated into the training set in order to obtain a model that is performing even better with the ICE-GENESIS data. A substantial improvement for water droplet recognition was required, and thus a supplementary number of 700 images of water droplets of all sizes was added to the training data set. The results of this second inspection are presented in Fig. 6b. The final testing results improved significantly: the overall accuracy rose from 60 % to 82 %, with 100 % recall for WD and 92 % accuracy. The results for the HPC class changed considerably as well. While before the assimilation many additional images would have been misidentified as HPCs, after the assimilation, even if the class-specific recall is lower (from 73 % to 52 %) the accuracy is higher (from 42 % to 88 %). This shows that the updated version of the model circumvents one of the difficulties humans are also subject to; namely, a heavily rimed plate and a spherical graupel particle are difficult to distinguish from OAP images, with the lack of surface information being one of the reasons. While it is acceptable to classify heavily rimed HPCs as CPs, the inverse classification is problematic for the study of microphysical processes. Overall, the above test improved the algorithm's predictions and its scientific interpretability.

2.3.2 PIP

Figure 7a describes the results of the class-specific testing. Co and CBC classes have acceptable recall and precision. The expected porosity between RAs, FAs, and CPs is confirmed. The porosity occurs at the expense of the FA class, which is an acceptable outcome. A result that has to be mentioned is the low accuracy of the HPC class (59%) and the inclusion of a high number of CPs in the HPC predictions. In order to try and rectify this confusion, initial testing data were assimilated and produced the results shown in Fig. 7b. The overall accuracy improved from 66% to 72%. Porosity between FAs and CBCs and between CPs and RAs is still considerable. A significant number of images of columns of only a few pixels are sometimes identified as FAs. Nevertheless these errors are acceptable because they do not harm the ability to make physical interpretations since they re-



Figure 3. (a) Confusion matrix for the CIP CNN. (b) Training report for the CIP CNN.



Figure 4. (a) Confusion matrix for the HVPS CNN. (b) Training report for the HVPS CNN.

spectively emanate from the same microphysical processes. Finally, the most problematic class in the original testing, HPCs, was fixed after data assimilation with an accuracy of 91 % and a recall of 72 %.

2.4 Computation of number and mass size distributions for different morphological classes

The four probes have pixel resolutions of 150, 100, 25, and 10 μ m, for the HVPS, PIP, CIP, and 2D-S, respectively. The 2D-S and HVPS have arrays of 128 photodiodes and the CIP and PIP of 64. Habit recognition requires the definition of minimum particle size this lower size threshold was set to 20 pixels for the HVPS, PIP, and CIP and 30 pixels for the 2D-S. This difference is motivated buy the high number of artifact particles that was obtained from 2D-S images

with the 20 pixel requirement. In addition, truncated images were excluded from the classification. The "entire-in" criterion provides the best shape recognition capabilities, ensures the quality of the extracted 2-D information, and it the original sampling volume computation formula, providing an accurate calculation of particle concentrations (Knollenberg, 1970). Nonetheless, this criterion reduces the number of analyzed images especially for size ranges nearing the length of the photodiode array, beyond which, only elongated particles oriented perpendicularly to the detection array are kept. Despite this artificial filtering, no upper size limit has been set for any of the probes used. In particular because this filtering heavily depends on particle shapes, the classification opens the possibility to even quantitatively explore this effect.

Since the classification tool is not yet integrated within the processing tools of OAP images at the LaMP (Leroy et al.,



Figure 5. (a) Confusion matrix for the "all-OAP" CNN. (b) Training report for the all-OAP CNN.



Figure 6. Results of specific evaluation for the 2D-S classes in number (100 images per class). The horizontal axis describes user classification and the vertical one the CNN algorithm results. (a) First evaluation. (b) After assimilation of the data from panel (a) and additional water droplet images.

2016), single images are extracted, tagged with timestamps and identification keys, so that they can be merged with routinely processed image features such as perimeter and surface area. After padding the raw images, they are processed by the classification tool, and each image is registered as one line in a table (pandas DataFrame) with its identification keys, timestamp, geometric features, and classification score for each class (the sum of which is normalized). The classification score is rounded so that it becomes a categorical binary array (one 1 and zeros). In addition, the aircraft data, in particular altitude, temperature, humidity measurements, and position, are resampled and concatenated in additional columns. Estimation of the sample volume is necessary in order to link classification results to quantitative concentration values. Computing the sample volume, in the existing code, is a dynamic operation that takes into account all recorded particles within the time frame of 1 s and the operational time of the probes. No direct method can therefore retrieve a particle by particle sample volume. In the future, the classification tool has to be integrated within the existing feature extraction IDL routines. For now, a simple approach was developed to retrieve class-specific number and mass size distributions, which easily translate to number and mass concentrations. The method is detailed below and illustrated in Fig. 8.



Figure 7. Results of specific evaluation for the PIP classes in number (100 images per class). The horizontal axis describes user classification and the vertical one the CNN algorithm results. (a) First evaluation. (b) After data assimilation.



Figure 8. Plots of morphology-specific size distributions from 2D-S data obtained in (a) Step 3, (b) Step 4, and (c) Step 5 of the computation of habit-specific particle size distribution (PSD). The data plotted correspond to the data gathered during flights 6, 7, and 8 of the ICE-GENESIS field campaign between +2 and -9 °C (see Sect. 3.1).

- 1. Particles are filtered, in time, in space (including altitude), or by relevant physical parameters such as temperature. The result of this step is a collection of particles gathered under user-defined filters.
- 2. A binned D_{max} is calculated for each particle. Particles are then grouped (summed) within bins corresponding to the bins used in the LaMP routine for particle size distribution (PSD) computation: the bin width is 10 µm (100 µm), and the first bin center is 5 µm (50 µm) for the 2D-S (PIP). The result is a pseudo class-specific PSD, uncorrected by the sample volume.
- 3. The number of images per class is normalized by the total number of particles in each bin, yielding a fraction of each type of particle in each bin (see Fig. 8a).
- 4. After applying the filters (used in step 1) to the PSD obtained through the LaMP routine, a PSD corresponding to the obtained class-specific normalized PSD is obtained. By multiplying the normalized class-specific pseudo PSD by its corresponding PSD bin by bin, the class-specific PSD is obtained (see Fig. 8b). The stacked PSD's envelope is the former PSD, obtained with the LaMP routine.
- 5. Plotting each morphology separately yields classspecific PSDs such as in Fig. 8c. And consequently, the sum of each PSD gives concentration values.

As a side note, the corresponding mass PSD can also be retrieved using the estimated mass of each particle, though Baker and Lawson (2006) combined single parameter parametrization or mass laws specific to each particle type. However, this step is not within the scope of this paper.

3 Application of the methodology to the ICE-GENESIS campaign

3.1 The ICE-GENESIS data set

The January 2021 ICE-GENESIS field campaign took place in the Swiss Jura mountains, over the Lachaux-de-Fond airport (Billault-Roux et al., 2023). The primary objective of this airborne experiment is to document snow conditions in the 0 to -8 °C temperature range, in terms of hydrometeor number concentrations, sizes, and shapes. Snowfall environments were found by sampling winter frontal clouds, whose precipitation were enhanced through the orographic effect of the mountainous terrain. OAPs were used to obtain the required measurements with high reliability. Of the most commonly used OAPs, four were mounted on the SAFIRE ATR-42, including a 2D-S, a 25 µm resolution CIP, a PIP, and an HVPS. The variety of resolution and size ranges of the probes provided measurements of hydrometeors in the 10 µm to 1.9 cm size range, with two extensive overlap regions between the 2D-S and CIP and between the PIP and HVPS. In addition, it allowed us to have adapted resolution for particles of various sizes, meaning that, given the CNN models presented in Sect. 2, morphological recognition of particles was possible from 300 μ m up to 19.2 cm. Results of the campaign documenting snow with these OAP measurements are presented in two conference papers (Jaffeux et al., 2023a, b).

The airborne RASTA W-band radar reflectivity vertical profiles for flights 6, 7, and 8 of the ICE-GENESIS campaign are presented in Fig. 9. Five cloud segments have been selected, corresponding to rather deep frontal clouds that were sampled continuously. For segment 3, a warm front was sampled and for the other segments cold fronts. Together, they represent 7 flight hours, where the cloud tops reached temperatures below -20 °C with bases nearing the melting point of water, meaning dendrites, plate-like, and columnar crystal types could be observed simultaneously with large snowflakes and rimed particles. In the present section, the data gathered within these periods are examined as a whole.

Figure 10 shows the flight distance corresponding to 1°C temperature intervals. About 25% of the total 3473 km were performed around -2°C. The dendritic and plate-like growth regions (between -20 to -12°C, and between -12 to -8°C, respectively) were not sampled, but some large precipitating crystals of these types were reported during the flights, and some smaller ones could be observed within larger aggregates. However, the columnar growth region (-8 to -5°C) was well explored. Supercooled liquid droplet pockets were marginally detected by the Cloud Droplet Probe outside the melting layer's vicinity.

The total PSDs were obtained with the all-in method and are presented in Fig. 11 for each OAP. With the exception of the ends of the spectra of both DMT instruments (CIP and PIP), where values tend to fall off before D_{max} reaches the photodiode array length, the four probes agree well over their overlapping range. In addition, the junction of both pairs of spectra is straightforward. In summary, all four spectra are very compatible.

For the 2D-S, only the horizontal channel was processed. In total, 410 882, 578 164, 986 965, and 257 103 images met the size requirements and all-in criterion for the 2D-S, CIP, PIP, and HVPS, respectively. They were analyzed by both specific and global CNN models. It can be noted that comparable numbers of usable images were gathered in just 7 flight hours of a single campaign with each of the four OAPs as in Przybylo et al. (2022), which obtained 970 000 images from 12 full campaigns using the CPI. This contrast highlights the limitations of the CPI and the advantages of using OAPs instead, even for morphological analysis. Using the methodology described in Sect. 2.4, the results of the shape recognition were combined with the total PSD presented in Fig. 11. The obtained PSDs are described in the next subsection. For each morphology-specific PSD, total concentration values were obtained by summing all the bins of each spectrum. These concentrations were then normalized to give the pie charts shown in Table 1. They reflect the type of ice parti-



Figure 9. RASTA reflectivity time series for each flights and identification of the five selected cloud segments, symbolized by the black rectangles and numbered accordingly.



Figure 10. Flight distance during legs as a function of temperature during the five selected flight segments shown in Fig. 9.

cles that were encountered within the data set for each probe and identified with the global and specific models presented in Sect. 2.2 and 2.3.

For the fine-resolution probes (2D-S and CIP), Co is the most frequent hydrometeor type, followed by CPs. For the coarser-resolution probes (PIP and HVPS), CPs, RAs, and FAs are the most frequent types. Both the specific and global models exhibit the same general trend: as pixel size gets larger, the relative concentration of Co decreases, and the relative concentration of FA increases. The specific and global models are consistent for the 2D-S, CIP, and HVPS, keeping in mind that the RA class is included in the CP one for the global model. Concerning the PIP, significantly fewer FAs are identified by the global model in comparison with the specific model. This discrepancy likely stems from different class definitions inferred from the corresponding training sets. This disparity shows that the definitions of the FA, CP, and RA classes differ between the PIP-specific and global models. The consistency of the classes across various instruments is, in this regard, a significant benefit of the global model.

3.2 Analysis of morphology-specific size distributions

In the present subsection, particle size distributions obtained with specific and global models are compared for each class or set of classes. The objectives are to assess the compatibility of the distributions obtained for each instrument and to evaluate which of the specific or global models yields the better results.

3.2.1 Compact particles and rimed aggregates

Images of quasi-spherical, densely rimed graupel-like particles make up the CP class. RA images are rimed in a similar way, but their shapes suggest underlying aggregates. This



Figure 11. Total particle size distributions for each OAP obtained with the ICE-GENESIS data set. Transparency and dots are used to distinguish the regions of each spectrum below the size threshold and above the photodiode array length.

leads to a significant porosity between these two categories. These two classes are designed to reveal the occurrence of riming and/or aggregation in clouds. The CP and RA categories of the HVPS and PIP may contain more particles than their 2D-S or CIP counterparts because unrimed aggregates cannot be distinguished from these particles at higher pixel resolutions. Therefore, the PIP and HVPS classes should likely overestimate the number of these particles compared to the 2D-S and CIP ones. Figure 12 presents the CP- and RA-specific particle size distributions for all four probes and both models.

The results obtained with the specific models for the RA class (Fig. 12a) show some consistency between the PIP and HVPS. The agreement between 2D-S and CIP for the CP class (Fig. 12b) is satisfactory, but the agreement between PIP and HVPS is significantly worse. At 3 mm, the HVPS curve is 1 order of magnitude above the PIP, and this difference only grows larger with increasing particle size. The curves corresponding to the 2D-S and CIP do not match well with the curves from the PIP and HVPS, the former beginning at values 1 order of magnitude below the highest CIP bin size. This could be due to the absence of a RA class for the two lower-resolution probes. The addition of the RA and CP class for the HVPS and PIP was performed and yielded the data shown in Fig. 12c. This operation only marginally improves the coherence of the two sets of curves, contradicting the earlier hypothesis that the HVPS and PIP should overestimate the number of these particles. However, there is one element of explanation left to investigate: the class definition for each model. Every model constructs an abstraction of what a CP or RA ought to be based on the training sets linked to the particular classes as well as in comparison to the other classes that have been defined. As can be seen in Fig. 12d, the possibility to fill in the size range gap for the CP class (that now includes the former RA class) improves significantly when a single model is utilized to identify the images from the various probes. However, the overlaps between the two pairs of OAP are slightly worse in comparison with the specific models.

3.2.2 Columns

Co consists of images of single columns or needles. This class is meant to identify deposition growth in the columnar regime. Throughout the training and testing phases, the precision of this class was remarkable for every model. However, elongated particles are not necessarily columns, and with coarse pixel resolution, misidentification of images is possible if the training set does not include enough elongated particles in other classes, such as FAs. The corresponding specific PSDs are presented in Fig. 13.

Figure 13a presents the Co-specific particle size distributions obtained with the specific models. These results are consistent across the 2D-S and CIP, but the PIP and HVPS disagree strongly, with values 1 order of magnitude above the PIP ones. The curves from the CIP or 2D-S match rela-

L. Jaffeux et al.: Morphology-dependent size distributions from in situ imagers

Table 1. Pie charts representing the hydrometeor fractions for total number concentration, by morphological class, and for each OAP identified with the global and specific models from the ICE-GENESIS data set.



tively well with the PIP or HVPS ones, and it is difficult to figure out whether the PIP or the HVPS matches best with the higher-resolution probes. Concerning the global model (see Fig. 13b), both the 2D-S-CIP and the PIP-HVPS pairs agree well on their respective overlap ranges. In addition, joining the highest CIP and lowest PIP size bins is straightforward.

3.2.3 Capped columns

The CC class features images of capped columns that may appear as "H"-shaped particles depending on their orientation. These particles are generally of sizes below 2 mm. Aggregates of a few columns can be mistaken for capped columns, especially with lower-resolution probes. The PSDs corresponding to this class are shown in Fig. 14.

For the specific models (Fig. 14a), only the 2D-S and CIP are available. Except for a few points, both distributions are very different, in particular for the smaller bins.

Below 500 µm, the 2D-S points are very high. The corresponding images show diffracted columns that look very similar to capped columns. Such diffraction patterns were already reported in Jaffeux et al. (2022). For the global model (Fig. 14b), both CIP and 2D-S distributions show similar shapes and agree above 600 µm. CC distributions are also now obtained for the PIP and HVPS. With respect to the 2D-S spectrum, much fewer diffracted columns are wrongly classified in the first bins. Important to note is the absence of HVPS images within the training set. Significantly more HVPS CCs are found for size bins below 5 mm compared to the PIP. The visual inspection of these images revealed pictures of aggregates of a few columns or fragile aggregates with an H shape, which can understandably be mistaken for capped columns, constituting legitimate identification errors. Similarly, for the PIP distribution, it is unlikely that 6 mm capped columns were encountered during the three



Figure 12. Total particle size distributions for each OAP obtained with the ICE-GENESIS data set for the CP and RA classes and identified with specific models (**a**, **b**, **c**) and the global model (**d**). The black points show the distribution points above the photodiode array length.



Figure 13. Total particle size distributions for each OAP obtained with the ICE-GENESIS data set for the Co class and identified with specific models (a) and the global model (b). The black points show the distribution points above the photodiode array length.

ICE-GENESIS flights. This implies that the rarely identified capped columns are particles whose images look like capped columns rather than real capped columns. However, the total concentrations associated with this class are significant for the coarse resolution instruments (3.44 % and 3.42 % of the mean concentrations for the HVPS and PIP, respectively – values from Table 1) above the CBC concentrations. For these two instruments, the CC class could be grouped with

the CBC or FA classes. Alternatively, the global training set could be enriched through the assimilation of some of the misclassified particles. This may result in increased precision and giving the classification algorithm the ability to better distinguish capped columns from other particle types for the PIP and HVPS.



Figure 14. Total particle size distributions for each OAP obtained with the ICE-GENESIS data set for the CC class and identified with specific models (a) and the global model (b). The black points show the distribution points above the photodiode array length.

3.2.4 Combinations of bullets and columns

Images of aggregates of columns and bullet rosettes can be found in the CBC class. Members of this class can only be classified by identifying individual columnar monomers in a larger image. Higher-resolution images provide a more precise view of individual columnar monomers, aiding in accurate classification. This means that the pixel resolution has a major influence on this class. CBCs that are detected by highresolution probes are typically aggregates of smaller columns than those that are imaged and detected at coarser resolution. Figure 15 shows the corresponding PSDs.

The results of the specific models are presented in Fig. 15a. For the 2D-S and CIP, plateaus are reached above 400 and 1000 μ m, respectively. Below 1 mm, the difference between the two curves is likely a consequence of the previously mentioned resolution effect, considering the 2D-S and CIP resolution are significantly different (10 and 25 μ m, respectively). The PIP bin concentration values are higher for the PIP over its optimal range compared to the HVPS, which can be similarly explained. However, the PIP and HVPS resolutions are relatively close (100 and 150 μ m, respectively). A possible explanation is the difference between the two definitions of the CBC class for each probe. The global model shows similar results, but it brings the PIP and HVPS much closer to one another in particular above 5 mm.

The results of the specific models are presented in Fig. 15a. For the 2D-S and CIP, plateaus are reached above 400 and 1000 μ m, respectively. Given that the 2D-S and CIP resolutions are significantly different (10 and 25 μ m, respectively), the difference between the two curves below 1 mm is likely due to the resolution effect discussed above. The PIP bin concentration values are higher for the PIP over its optimal range compared to the HVPS, which can be similarly explained. However, the PIP and HVPS resolutions are relatively close (100 and 150 μ m, respectively). A possible explanation is the difference between the two definitions of the CBC class for

each probe. Comparable results are obtained with the global model for the 2D-S, CIP, and HVPS. However, the values of the PIP distribution decrease by a factor of 2, bringing it much closer to the HVPS distribution, in particular above 5 mm. This suggests that the differences between the PIP and HVPS using specific models may be due to variations in how the CBC class is defined for each probe.

3.2.5 Complex assemblages

CAs constitute a class of ice particles whose images show sharp edges and transparency and frequently display multiple plates or sector plates. The corresponding particles indicate the occurrence of deposition in highly saturated environments and possibly the aggregation of plates and dendrites. The high level of detail that is required for their identification was only found in 2D-S images. The PSDs obtained for this class are shown in Fig. 16.

Out of the four specific models, only the one trained for the 2D-S possesses this class (see Fig. 16a). For the global model (Fig. 16b), this class appears for each probe. Concerning the 2D-S global model output, it corresponds to the specific model almost point by point. For the CIP, a similar curve is produced but with much lower concentration values and at larger bin sizes. The inspection of the corresponding images is surprising as they share a common feature with the 2D-S defined class: the sharpness of their edges. The CA class was therefore neither discarded nor assimilated into another class in the global model. For PIP and HVPS, the distributions are bell-shaped curves centered at 5 and 7 mm, respectively. These curves exhibit a staggered pattern comparable to the CBC curves. This implies that pixel resolution has a similar effect on the CA class. For probes other than the 2D-S, the CA class is, however, very marginally seen within the present data set for probes other than the 2D-S, accounting for 2.98 %, 0.49 %, 0.28 %, and 0.15 % for the 2D-S, CIP, PIP, and HVPS total number concentration, respectively. Significant uncertainty is associated with the identi-



Figure 15. Total particle size distributions for each OAP obtained with the ICE-GENESIS data set for the CBC class and identified with specific models (a) and the global model (b). The black points show the distribution points above the photodiode array length.



Figure 16. Total particle size distributions for each OAP obtained with the ICE-GENESIS data set for the CA class and identified with specific models (a) and the global model (b). The black points show the distribution points above the photodiode array length.

fication of CA particles with probes other than the 2D-S. However, the shapes of the curves obtained are encouraging with respect to the potential to identify particle morphology in poorly resolved images beyond human capabilities. In this instance, a model trained on high resolution images was used on lower-resolution images, yielding satisfactory results. These findings suggest that the identification of CA particles using probes other than the 2D-S may be feasible with further research.

3.2.6 Hexagonal planar crystals

HPCs are the class of single plates and dendrites. The 6-fold symmetry and the planarity of the corresponding particles are the two defining features of this class. Learning these characteristics was found to be particularly challenging and showed relatively low precision or recall during the testing for every model. As a reminder, recall was low for 2D-S and CIP, precision was low for the HVPS, and both were low for the PIP. With respect to its dependency on pixel resolution, the HPC is a special case. A minimum size greater than the minimum threshold for each probe might be needed to accurately identify them because of the effect of pixelization on sharp edges of a few pixels in length. However, with an adequate pixel number, resolution should have no impact. The PSDs obtained for the HPC class are shown in Fig. 17.

For the specific models (see Fig. 17a), a large difference is noted between the 2D-S and CIP below 1 mm, with the CIP being about 1 order of magnitude above the 2D-S. Despite their divergence at large bin sizes, the PIP and HVPS curves are generally close, with the PIP values being higher than the HVPS. This discrepancy can be attributed to the high recall and low precision on the HVPS testing set, indicating a possible overestimation on the HVPS side. For the global model (Fig. 17b), the CIP and 2D-S curves are much closer, with the few first CIP values decreasing and those of the 2D-S increasing slightly in comparison with the specific models. The HVPS and PIP curves are also relatively closer to one another. Overall, the results suggest that the global model may provide a more accurate identification for the HPC class compared to the specific models. However, the visual inspec-



Figure 17. Total particle size distributions for each OAP obtained with the ICE-GENESIS data set for the HPC class and identified with specific models (a) and the global model (b). The black points show the distribution points above the photodiode array length.

tion of the classified images shows that improvements are still desired to reduce the number of particles that were misclassified in this class. Further data assimilation could potentially enhance the accuracy of the global model for classifying HPC particles.

3.2.7 Fragile aggregates

Weakly linked aggregates with monomers that cannot be recognized as specific crystal types make up the FA class. Because it is likely to contain CBC or CA particles, whose details could not be obtained for low-resolution probes, this class depends on pixel size. Figure 18 displays the PSDs that were obtained for the FA class.

For the specific models (Fig. 18a), the 2D-S and CIP agree for bin sizes above 500 µm to some extent. Similarly to the results obtained for CPs with the particular models, the PIP and HVPS curves show a significant divergence. The 2D-S-CIP and HVPS-PIP distributions are difficult to join, in agreement with the previously mentioned resolution dependency of this class. The classification with the global model (Fig. 18b) considerably alters the shape of the CIP distribution, with a minimum at 1 mm. The 2D-S size distribution is only significantly lower in the first bins compared to the specific curve. Finally, the PIP and HVPS curves are closer using the global model. This result is reassuring considering the similar pixel resolution between both instruments. For the FA class, which is more loosely defined compared to the other class, many of the slight differences between the different morphology-specific PSDs might have implications for the FA PSD.

3.2.8 Water droplets

The WD class is composed of very smooth, spherical images that can be attributed to water droplets. These classes were originally defined for the 2D-S and CIP only because of the difficulty of finding liquid or frozen droplets above 2 mm within the data available to pick from. This class was well trained for all the models. It can be noted that finding water droplets in OAP images can be particularly useful to study mixed-phase clouds. The current method is restricted to sizes greater than a certain pixel threshold, but it should perform better than more established approaches that distinguish water droplets using the circularity parameter. The obtained PSDs are presented in Fig. 19.

The 2D-S results for the specific and global models are remarkably similar (Figs. 19a and 19b). However, because there were too few big water droplets in the CIP training set, the CIP-specific model failed to detect larger droplets, whereas the global model could. In addition, it could find WD in the PIP data but not in the HVPS. The differences in performance between the specific and global models highlight the importance of having a diverse training data set. By incorporating data from multiple sources, the global model was able to better generalize and detect larger water droplets across different instruments.

3.2.9 Artifact classes: fragmented and diffracted particles

Two artifact classes were originally defined for the 2D-S (Dif) and the HVPS (FP). Their purpose is to remove artifacts that remain after routine artifact treatments. In the case of diffracted images, these are very specific to the 2D-S. The fragmented particles that were found in the HVPS data are significantly different compared to what is usually denoted as shattering in the literature (Korolev and Isaac, 2005). They are not the very small particles that appear in high amounts in the size distribution and that do not affect the current methodology since a relatively high size threshold is used. Instead, this class refers to rather large particles which appear as a "cloud of particles" on a single image. The corresponding size distributions are shown in Fig. 20.



Figure 18. Total particle size distributions for each OAP obtained with the ICE-GENESIS data set for the FA class and identified with specific models (a) and the global model (b). The black points show the distribution points above the photodiode array length.



Figure 19. Total particle size distributions for each OAP obtained with the ICE-GENESIS data set for the WD class and identified with specific models (a) and the global model (b). The black points show the distribution points above the photodiode array length.

The Dif PSD obtained with the 2D-S specific model is plotted in Fig. 20a and exhibits the same shape as the PSD resulting from the use of the global model on the same data (see 2D-S curve in Fig. 20b). However, the values are much lower in the case of the global model, which shows in the total concentration fraction varying from 7.59% for the specific model to 4.73 % for the global model (see Table 1). The application of the global model to the data gathered with other probes yielded very few images in the Dif class, resulting in 0.78%, 0.06%, and 0.05% of the total concentration for the CIP, PIP, and HVPS, respectively. For the FP class, the specific model for the HVPS identifies more FPs compared to the global model, with 1.79 % against 1.28 %. However, the corresponding curves exhibit a similar shape once more (see HVPS PSD in Fig. 20c and d). The global model identified a significant number of FPs for the CIP that amounts to 3.23 % of the total measured concentration for this instrument. A visual examination of the corresponding pictures demonstrated that these identifications were, in fact, accurate. The singularly high number of FPs within the CIP compared to the other instruments may be attributed to differences in optics and electronics and the design of its arms with respect to the position of its laser beam. Compared to the probe with the most similar optics and electronics, the arms of the PIP extend outwards, whereas the arms of the CIP are parallel, which could produce more of these shattering events. In any case, the FP PSD deserves further investigation for the CIP.

4 Conclusions

In the present article, morphological data from OAPs were used in a new way. First, three new CNNs were trained, two of which were specific models for the CIP and HVPS, and the last one of which was a global model that can be used on all OAPs. Two CNNs that were previously developed for the 2D-S and the PIP were tested and improved. A methodology was presented to obtain size distributions specific to morphology by combining particle size distributions with CNN models for particle identification.



Figure 20. Total particle size distributions for each OAP obtained with the ICE-GENESIS data set for the Dif (a, b) and FP (c, d) classes and identified with specific models (a, c) and the global model (b, d). The black points show the distribution points above the photodiode array length.

Then, the ICE-GENESIS data set, which comprises all four OAPs, was presented. The morphology-specific PSDs were described and discussed at a high level of detail for each probe and for both specific and global models. For compact particles, columns, combinations of bullets and columns, hexagonal planar crystals, and water droplets, the global model had the advantage of unifying the results obtained from the different instruments compared to the specific models. For capped columns and complex assemblages, the global model was able to extrapolate, with relative success, the classes of the higher-resolution probes to the coarserresolution ones. However, major uncertainty about the accuracy of these classes remains when used on PIP or HVPS images, especially for capped columns. For the present data set, these two classes are of minor importance in the HVPS and PIP size range. For the artifact classes, namely diffracted and fragmented particles, using the global model decreased the number of images identified for the specific probes for which these classes were originally defined. However, it revealed a high number of fragmented particles within the CIP images. This constitutes, in itself, an important result of the study. Further investigations on other data sets will help to determine the limitations and strengths of the global model in classifying different types of particles in OAP images.

The study demonstrated the effectiveness of utilizing CNN models for morphological data interpretation and, in particular, the advantages of using a single global CNN model. The obtained morphology-specific particle size distributions represent a major step in the characterization of atmospheric ice and water particles. A few areas of improvement could be identified in the process of examining these spectra. With the training methodology now perfected, these can be rectified through data assimilation, further enhancing the accuracy and reliability of the global CNN model in analyzing OAP data.

The developed models can be applied to decades of OAP data. Since they were developed, these instruments have been the cornerstone of in situ aircraft measurements of clouds and the gold standard for cloud model outputs. By reducing 50 years of hydrometeor observation across the world to their most pertinent features (size, shape, and concentration), significant improvements in the current understanding of cloud processes can likely be achieved. However, putting morphology at the center of ice cloud observation requires rethinking how ice clouds are conceptualized and modeled.

The differentiation of cloud and precipitation particles into several interacting populations brings forward the complexity of the interactions between different hydrometeor types. During collisions between two hydrometeors, breakup, riming, or aggregation may happen depending on their respective types and sizes. These are important events that trigger secondary ice production and influence precipitation rates. This approach therefore has the potential to improve the understanding of precipitation patterns and cloud life cycles, for example.

Convolutional neural networks are semi-supervised learning algorithms that deduce image features that can characterize individual morphological classes from manually sorted data. Under the condition that the training is successful, the obtained classification model can be considered an objective class definition tool that matches the human ability beyond any worded description. The presented models are efficient and were among the first to be developed for optical array probes. For these reasons, the training sets gathered at the LaMP may be considered the first reference data sets after manual inspections and validation by other ice microphysics specialists from different research facilities around the globe. The data set can of course be expanded and evaluated regularly. This will allow for harmonized, consistent, and comparable results across different research teams utilizing automatic classification tools. Finally, sharing the training data will promote collaboration and exchanges in the fields of ice microphysics and cloud research. With the recent advances in artificial intelligence, the authors believe that collaboration and transparency in data sharing should be promoted and that, collectively, scientists can create the next generation of analysis tools. The present contribution has already shown that such developments are possible in the field of cloud microphysics with ice particle images.

Code and data availability. Training and testing data, as well as trained CNN models, are publicly available at https://github.com/LJaffeux/JAFFEUX_et_al_AMT_2024 (last access: 24 May 2025, labeled binary images; https://doi.org/10.5281/zenodo.15573218, Jaffeux, 2025). Developed Python scripts can be provided upon request to the authors.

Author contributions. LJ wrote the Python code, compiled the data, trained the CNNs, and wrote the article. JB performed preliminary data labeling for the HVPS probe. AS and PC revised the manuscript and helped during the reviewing process.

Competing interests. The contact author has declared that none of the authors has any competing interests.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, pub-

lished maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors.

Acknowledgements. We acknowledge Wiebke Frey for her dedication and caring attention as the editor of this article. We also extend our gratitude to the reviewers for their constructive feedback and valuable insights. Data used in this study (from past projects EUREC4A, EXAEDRE, ICE GENESIS, and HAIC/HIWC) were mainly obtained using ATR-42 and Falcon 20 research aircraft managed by SAFIRE (French facility for airborne research), thereby operating the cloud in situ instrumental payload from the French INSU/CNRS Airborne Measurement Platform (PMA). Minor data were contributed from the past AFLUX project, which is part of the German Transregional Collaborative Research Centre TR 172 (ArctiC Amplification: Climate Relevant Atmospheric and SurfaCe Processes, and Feedback Mechanisms (AC)3).

Review statement. This paper was edited by Wiebke Frey and reviewed by two anonymous referees.

References

- Baker, B. and Lawson, R. P.: Improvement in Determination of Ice Water Content from Two-Dimensional Particle Imagery. Part I: Image-to-Mass Relationships, J. Appl. Meteorol. Clim., 45, 1282–1290, https://doi.org/10.1175/JAM2398.1, 2006.
- Billault-Roux, A. C., Grazioli, J., Delanoë, J., Jorquera, S., Pauwels, N., Viltard, N., Martini, A., Mariage, V., Le Gac, C., Caudoux, C., Aubry, C., Bertrand, F., Schwarzenboeck, A., Jaffeux, L., Coutris, P., Febvre, G., Pichon, J. M., Dezitter, F., Gehring, J., Untersee, A., Calas, C., Figueras i Ventura, J., Vie, B., Peyrat, A., Curat, V., Rebouissoux, S., and Berne, A.: ICE GENESIS: Synergetic aircraft and ground-based remote sensing and in situ measurements of snowfall microphysical properties, B. Am. Meteor. Soc., 104, E367–E388, 2023.
- Duroure, C.: Une nouvelle méthode de traitement des images d'hydrométéores données par les sondes bidimensionnelles, Journal de recherches atmosphériques, https://hal.uca.fr/ hal-01950254 (last access: 24 May 2025), 1982.
- Emersic, C. and Saunders, C.: Further laboratory investigations into the relative diffusional growth rate theory of thunderstorm electrification, Atmos. Res., 98, 327–340, 2010.
- Field, P., Heymsfield, A., and Bansemer, A.: Shattering and particle interarrival times measured by optical array probes in ice clouds, J. Atmos. Ocean. Tech., 23, 1357–1371, 2006.
- Fukuta, N. and Takahashi, T.: The Growth of Atmospheric Ice Crystals: A Summary of Findings in Vertical Supercooled Cloud Tunnel Studies, J. Atmos. Sci., 56, 1963–1979, https://doi.org/10.1175/1520-0469(1999)056<1963:TGOAIC>2.0.CO;2, 1999.
- Jaffeux, L.: LJaffeux/JAFFEUX_et_al_AMT_2024: Zenodo, Zenodo [data set and code], https://doi.org/10.5281/zenodo. 15573218, 2025.

L. Jaffeux et al.: Morphology-dependent size distributions from in situ imagers

- Jaffeux, L., Schwarzenböck, A., Coutris, P., and Duroure, C.: Ice crystal images from optical array probes: classification with convolutional neural networks, Atmos. Meas. Tech., 15, 5141–5157, https://doi.org/10.5194/amt-15-5141-2022, 2022.
- Jaffeux, L., Coutris, P., Schwarzenboeck, A., and Dezitter, F.: Snow Particle Characterization. Part B: Morphology Dependent Study of Snow Crystal 3D Properties Using a Convolutional Neural Network (CNN), SAE Technical Paper, SAE International, https://doi.org/10.4271/2023-01-1486, 2023a.
- Jaffeux, L., Schwarzenboeck, A., Coutris, P., Febvre, G., Dezitter, F., Aguilar, B., Billault-Roux, A. C., Grazioli, J., Berne, A., Köbschall, K., Jorquera, S., and Delanoe, J.: Snow Particle Characterization. Part A: Statistics of Microphysical Properties of Snow Crystal Populations from Recent Observations Performed during the ICE GENESIS Project, SAE Technical Paper, SAE International, https://doi.org/10.4271/2023-01-1492, 2023b.
- Jensen, A. A. and Harrington, J. Y.: Modeling Ice Crystal Aspect Ratio Evolution during Riming: A Single-Particle Growth Model, J. Atmos. Sci., 72, 2569–2590, https://doi.org/10.1175/JAS-D-14-0297.1, 2015.
- Knight, C. A.: Observations of the morphology of melting snow, J. Atmos. Sci., 36, 1123–1130, 1979.
- Knollenberg, R. G.: The Optical Array: An Alternative to Scattering or Extinction for Airborne Particle Size Determination, J. Appl. Meteorol., 9, 86–103, https://doi.org/10.1175/1520-0450(1970)009<0086:TOAAAT>2.0.CO;2, 1970.
- Knorr, W., Dentener, F., Lamarque, J.-F., Jiang, L., and Arneth, A.: Wildfire air pollution hazard during the 21st century, Atmos. Chem. Phys., 17, 9223–9236, https://doi.org/10.5194/acp-17-9223-2017, 2017.
- Korolev, A. and Isaac, G. A.: Shattering during sampling by OAPs and HVPS. Part I: Snow particles, J. Atmos. Ocean. Tech., 22, 528–542, 2005.
- Krizhevsky, A., Sutskever, I., and Hinton, G. E.: Imagenet classification with deep convolutional neural networks, Adv. Neur. In., 25, 1097–1105, 2012.
- Leroy, D., Fontaine, E., Schwarzenboeck, A., and Strapp, J. W.: Ice Crystal Sizes in High Ice Water Content Clouds. Part I: On the Computation of Median Mass Diameter from In Situ Measurements, J. Atmos. Ocean. Tech., 33, 2461–2476, https://doi.org/10.1175/JTECH-D-15-0151.1, 2016.
- Locatelli, J. D. and Hobbs, P. V.: Fall speeds and masses of solid precipitation particles, J. Geophys. Res., 79, 2185–2197, 1974.
- Matsuo, T. and Sasyo, Y.: Melting of snowflakes below freezing level in the atmosphere, J. Meteorol. Soc. Jpn. Ser. II, 59, 10–25, 1981.
- McFarquhar, G. M., Timlin, M. S., Rauber, R. M., Jewett, B. F., Grim, J. A., and Jorgensen, D. P.: Vertical variability of cloud hydrometeors in the stratiform region of mesoscale convective systems and bow echoes, Mon. Weather Rev., 135, 3405–3428, 2007.
- Pasquier, J. T., Henneberger, J., Korolev, A., Ramelli, F., Wieder, J., Lauber, A., Li, G., David, R. O., Carlsen, T., Gierens, R., Maturilli, M., and Lohmann, U.: Understanding the history of two complex ice crystal habits deduced from a holographic imager, Geophys. Res. Lett., 50, e2022GL100247, https://doi.org/10.1029/2022GL100247, 2023.
- Praz, C., Roulet, Y.-A., and Berne, A.: Solid hydrometeor classification and riming degree estimation from pictures collected with a

Multi-Angle Snowflake Camera, Atmos. Meas. Tech., 10, 1335–1357, https://doi.org/10.5194/amt-10-1335-2017, 2017.

- Pruppacher, H. and Klett, J.: Microphysics of Clouds and Precipitation, Taylor & Francis, vol. 18, 116–119, https://doi.org/10.1007/978-0-306-48100-0, 2010.
- Przybylo, V., Sulia, K. J., Lebo, Z. J., and Schmitt, C.: Automated Classification of Cloud Particle Imagery through the Use of Convolutional Neural Networks, in: 101st American Meteorological Society Annual Meeting, AMS, online, 10–15 January 2021, https://ui.adsabs.harvard.edu/abs/2021AMS...10177736P/ abstract (last access: 24 May 2025), 2021.
- Przybylo, V. M., Sulia, K. J., Schmitt, C. G., and Lebo, Z. J.: Classification of cloud particle imagery from aircraft platforms using convolutional neural networks, J. Atmos. Ocean. Tech., 39, 405– 424, 2022.
- Rahman, M. M., Quincy, E. A., Jacquot, R. G., and Magee, M. J.: Feature Extraction and Selection for Pattern Recognition of Two-Dimensional Hydrometeor Images, J. Appl. Meteorol., 20, 521–535, https://doi.org/10.1175/1520-0450(1981)020<0521:FEASFP>2.0.CO;2, 1981.
- Schmitt, C. G., Järvinen, E., Schnaiter, M., Vas, D., Hartl, L., Wong, T., and Stuefer, M.: Classification of ice particle shapes using machine learning on forward light scattering images, Artificial Intelligence for the Earth Systems, 3, 230091, https://doi.org/10.1175/AIES-D-23-0091.1, 2024.
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R.: Dropout: A Simple Way to Prevent Neural Networks from Overfitting, J. Mach. Learn. Res., 15, 1929– 1958, http://jmlr.org/papers/v15/srivastava14a.html (last access: 20 May 2025), 2014.
- Vaillant de Guélis, T., Schwarzenböck, A., Shcherbakov, V., Gourbeyre, C., Laurent, B., Dupuy, R., Coutris, P., and Duroure, C.: Study of the diffraction pattern of cloud particles and the respective responses of optical array probes, Atmos. Meas. Tech., 12, 2513–2529, https://doi.org/10.5194/amt-12-2513-2019, 2019.
- Vázquez-Martín, S., Kuhn, T., and Eliasson, S.: Shape dependence of snow crystal fall speed, Atmos. Chem. Phys., 21, 7545–7565, https://doi.org/10.5194/acp-21-7545-2021, 2021.
- Westbrook, C. D., Hogan, R. J., and Illingworth, A. J.: The capacitance of pristine ice crystals and aggregate snowflakes, J. Atmos. Sci., 65, 206–219, 2008.
- Wu, Z., Liu, S., Zhao, D., Yang, L., Xu, Z., Yang, Z., Zhou, W., He, H., Huang, M., Liu, D., Li, R., and Ding, D.: Neural network classification of ice-crystal images observed by an airborne cloud imaging probe, Atmos.-Ocean, 58, 303–315, 2020.
- Wyser, K.: Ice crystal habits and solar radiation, Tellus A, 51, 937– 950, 1999.
- Zhang, R., Xiao, H., Gao, Y., Su, H., Li, D., Wei, L., Li, J., and Li, H.: Shape Classification of Cloud Particles Recorded by the 2D-S Imaging Probe Using a Convolutional Neural Network, J. Meteorol. Res., 37, 521–535, 2023.
- Zhu, S., Guo, X., Lu, G., and Guo, L.: Ice crystal habits and growth processes in stratiform clouds with embedded convection examined through aircraft observation in northern China, J. Atmos. Sci., 72, 2011–2032, 2015.