Parameterization of the collision-coalescence process using series of basis functions: *COLNETv1.0.0* model development using a machine learning approach

Camilo Fernando Rodríguez-Genó[1], Léster Alfonso[2]

[1]Atmospheric Sciences Centre, National Autonomous University of Mexico, Mexico City, 04510, Mexico
[2]Autonomous University of Mexico City, Mexico City, 09790, Mexico

Correspondence to: Camilo Fernando Rodríguez Genó (camilo.rodriguez@atmosfera.unam.mx)

Abstract. A parameterization for the collision-coalescence process is presented, based on the methodology of basis functions. The whole drop spectra is depicted as a linear combination of two lognormal distribution functions, in which all distribution parameters are formulated by means of six distribution moments included in a system of equations, thus eliminating the need of fixing any parameters. This basis functions parameterization avoids the classification of drops in artificial categories such as cloud water (cloud droplets) or rain water (raindrops). The total moment tendencies are calculated using a machine learning approach, in which one deep neural network was trained for each of the total moment orders involved. The neural networks were trained using randomly generated data following a uniform distribution, over a wide range of parameters employed by the parameterization. An analysis of the predicted total moment errors was performed, aimed to stablish the accuracy of the parameterization at reproducing the integrated distribution moments representative of physical variables. The applied machine learning approach shows a good accuracy level when compared to the output of an explicit collision-coalescence model.

Keywords: cloud microphysics; collision-coalescence; lognormal distribution; microphysics parameterization; numerical modelling; machine learning; neural networks.

1 Introduction

Drop populations are well represented using drop size distributions (DSD). The first attempt at characterizing drop spectra using this approach was developed by Marshall and Palmer, whom employed exponential distributions based on drop diameter to describe the DSDs (Marshall and Palmer, 1948). Also, the use of a three-parameter gamma distribution have shown a good agreement with observations (Ulbrich, 1983). However, lognormal distributions have shown a better squared-error fit to measurements of rain DSDs than gamma or exponential distributions (Feingold and Levin, 1986; Pruppacher and Klett, 2010). For spherical particles such as cloud drops, a transformation of the DSD leads to a self-preserving form. This convergence is very fast for particles under a high Knudsen number regime (in the order of microseconds), when assuming high concentration values in the initial stages of the process. However, it has been confirmed that such a behaviour is not observed in low Knudsen

30 number environments (Dekkers and Friedlander, 2002; Lee et al., 1997). The analysis of several important characteristics of the Brownian coagulation process showed that the lognormal distribution adequately represents the particle distributions. Further, several results indicate that the required time to attain a self-preserving form of a particle spectra is related to the initial geometric standard deviation as well (Lee et al., 1984, 1997). In addition, some authors have employed this type of distribution function to parameterize cloud processes with promising results (Clark, 1976; Feingold et al., 1998; Huang, 2014).

35 There are two main approaches to modelling cloud processes: the explicit approach (bin microphysics) and the bulk approach (bulk microphysics). Bin microphysics is based on the discretization of a droplet size distribution (DSD) into sections (bins), and calculates the evolution of the DSD due to the influence of different processes that could be dynamical and/or microphysical (Berry, 1967; Berry and Reinhardt, 1974; Bott, 1998a; Khain et al., 2004, 2010). The core of this method is the solution of the Kinetic Coagulation Equation (KCE) (von Smoluchowski, 1916a, 1916b) for the collision-coalescence of liquid

40 drops, (also known as Stochastic Coalescence Equation or Kinetic Collection Equation within the cloud physics community), in a previously designed grid, that could be over mass or radius. Thus, previous knowledge of the characteristics or parameters of the distributions is not necessary. This way of solving the KCE is very accurate, but its operational utility is compromised because it is computationally very expensive, due to the need to calculate a huge amount of equations, which number ranges from several dozens to hundreds, at each grid point and time step.

45 In the case of bulk microphysics, the KCE is parameterized and the evolution of a chosen set of statistical moments related to physical variables is calculated, instead of the evolution of the DSD itself. A pioneer approach to this kind of parameterizations can be found at (Kessler, 1969), where it is introduced a simple but accurate representation of the autoconversion process. One or two-moment parameterizations are common (Cohard and Pinty, 2000; Lim and Hong, 2010; Milbrandt and McTaggart-Cowan, 2010; Morrison et al., 2009; Thompson et al., 2008), however, recently it has been extended to three-moment

50 parameterizations such as (Huang, 2014; Milbrandt and Yau, 2005). This type of parameterization is computationally efficient, which make it popular within the operational weather forecasting community. The main disadvantage of this approach is that the equations for solving the rates of the p-th moment include moments of a higher order, so the system of equations employed to calculate the evolution of the moments is not closed (Seifert and Beheng, 2001). This could be avoided by using predefined parameters for the distributions that describe the DSD, which normally take the form of exponential (Marshall and Palmer,

55 1948), gamma (Milbrandt and McTaggart-Cowan, 2010; Milbrandt and Yau, 2005) or lognormal distributions (Huang, 2014). Besides, artificial categories are often used to separate water substance (cloud and rain water), dependent of the drop radius, being $20 \, \mu m$ and $41 \, \mu m$ very popular thresholds (Cohard and Pinty, 2000; Khairoutdinov and Kogan, 2000), with the moments for each category being calculated by means of partial integration of the KCE.

An additional approach on this matter is the particle-based one, which is based on the application of a stochastic model such

60 as the Monte Carlo method to the coagulation (coalescence) of drop particles inside a cloud. This method have been approached from a number of perspectives. For example (Alfonso et al., 2008) analysed the possible ways of solving the KCE by using a Monte Carlo algorithm and several collision kernels, with good correspondence between the analytical and numerical approaches for all the kernels, by estimating the KCE following the Gillespie's Monte Carlo algorithm (Gillespie, 1972) and

several analytical solutions. Also, the possible implications of this approach for cloud physics are discussed. Other variants of
65   this approach are analysed in (Alfonso et al., 2011), and more it has been used to simulate the subprocesses of autoconversion
and accretion applying a Monte Carlo-based algorithm within the framework of Lagrangian cloud models (Noh et al., 2018).
This approach is accurate, and represents well the stochastic nature of the collision-coalescence of drops, but it is very
expensive computationally.

An alternative to these main approaches can be applied by using a hybrid approach to parameterize the cloud microphysical
70   processes. This approach simulates the explicit approach in the way that it describes the shape of the DSD through a linear
combination of basis functions (Clark, 1976; Clark and Hall, 1983), and it could be considered a middle point between bulk
and bin microphysics. A system of prognostic equations is solved to obtain the tendencies of the parameters of the statistical
distribution functions based on the evolution of its total moments (the combination of the statistical moments of the same order
of all distribution functions involved), describing their tendencies due to condensation and collision-coalescence. Since the
75   integration process covers the entire domain, the artificial separation of the droplet spectrum is avoided, making the terms
cloud droplet and rain drop meaningless (there is just drops), and it is possible to solve a fully closed system of equations
without the need to keep any parameter of the distribution constant. However, this integration can be made only once for all
parameters at each time step. Another advantage of this approach is its independence from a specific collision kernel type, as
is common in the bulk approach, that in order to obtain analytical expressions from the integrals of the KCE, a polynomial
80   type kernel such as the one derived by (Long, 1974) is frequently used. Having said that, a limitation of this approach is that
the total moment tendencies have to be solved at each time step for the needed parameters, or being previously calculated by
including a sufficiently wide range of parameters that should be stored in lookup tables that must be consulted several times
at every time step.

In order to eliminate the need to solve the rate equations for the total moments of the KCE at every time step (Thompson,
85   1968), or resort to the use of lookup tables, a Machine Learning (ML) approach is proposed to calculate the total moment
tendencies within this parameterization. With this approach, the use of one trained Deep Neural Network (DNN) for each
calculated total moment tendency will accelerate the calculations without sacrificing precision. Thus, the objective of this
study is to apply DNN to the parameterized formulation of the collision-coalescence process developed at (Clark, 1976) in
order to replicate the rate equations for the total moments, eliminating the need of memory expensive lookup tables or
90   processing-time expensive numerical solution of integrals.

Machine Learning is the study of computer algorithms that improve automatically through experience and by the use of data
(training) (Mitchell, 1997). Machine Learning algorithms build a model based on sample data in order to make predictions or
decisions without being explicitly programmed to do so (Koza et al., 1996). Machine learning algorithms are used in a wide
variety of applications, such as in medicine, email filtering, and computer vision, where it is difficult or unfeasible to develop
95   conventional algorithms to perform the needed tasks. In particular, neural networks (NN) are specially well suited for solving
non-linear fitting problems and for stablishing relationships within complex data such as the outputs of the KCE. In the field
of atmospheric sciences, the use of DNNs has been extended to the parameterization of subgrid processes in climate models

(Brenowitz and Bretherton, 2018; Rasp et al., 2018), while in cloud microphysics, the autoconversion process was parameterized using DNNs with a superior level of accuracy when compared with equivalent bulk models (Alfonso and Zamora, 2021; Sobhani et al., 2018). Also, a partial parameterization of collision-coalescence was tested in (Seifert and Rasp, 2020), which developed a ML parameterization that includes the processes of autoconversion and accretion, describing the droplet spectra as a gamma distribution, and establishing a comparative study that exposed the advantages and disadvantages of the use of ML techniques on cloud microphysics.

The research article is structured as follows: In section 2, the parameterization framework is described; In section 3, the procedures of DNN methodology are explained and the network architecture is introduced, the training data set is generated, and the DNN is trained and validated; section 4 includes a description of the explicit model used as reference solution; In section 5, the experiment design is explained; In section 6, the results are discussed, an assessment of the results is made by contrasting them with the reference solution, and the predicted total moment errors are analyzed; and in section 7 several conclusions are drawn.

2 Description of the collision-coalescence parameterization

Under the framework of this parameterization, any given drop spectrum can be approximated by a series of basis functions. Therefore, the distribution that characterizes the evolution of the spectrum is given by a linear combination of probability density functions as shown below:

$$f\langle r\rangle = \sum_{i=1}^{I} f_i\langle r\rangle \qquad (1)$$

where $f_i\langle r\rangle$ are the individual members of the set of distributions considered, and I stands for the number of distributions functions that make up the set. In the case at hand, a set of two statistical distribution is employed. At each time step, the rates of the parameters of each distribution will be calculated. It should be noted that, in any set of distributions considered, all its members will be of the same type of distribution. For the proposed parameterization, as described in (Clark, 1976), a distribution of log-normal type is used, as follows

$$\boldsymbol{f\langle r\rangle} = \frac{\boldsymbol{N}}{\sqrt{2\pi}\sigma r}\, \boldsymbol{e}^{[-(\ln r - \mu)^2/(2\sigma^2)]} \qquad (2)$$

Where μ and σ^2 stand for the mean and variance of $\ln r$ respectively. Considering that moment of order p of any distribution can be defined as (Straka, 2009)

$$N\overline{R^p} = \int_0^\infty r^p f(r)\, dr \qquad (3)$$

the following analytical solution of eq. (3) can be found for the moments of the lognormal distribution

125

$$\overline{R^p} = e^{p\mu + \frac{1}{2}p^2\sigma^2} \qquad (4)$$

Combining eqs. (1), (3) and (4), the p-th total moment of a linear combination of lognormal distributions could be expressed as (Clark and Hall, 1983)

$$N\overline{R^p} = \sum_{i=1}^{I} N_i \overline{R_i^p} = \sum_{i=1}^{I} N_i e^{p\mu_i + \frac{1}{2}p^2\sigma_i^2} \qquad (5)$$

Where the index i indicates each of the individual members of the set. Deriving eq. (5) with respect to time, we obtain the

130   tendencies of the total moments of a series of log-normal distributions

$$\frac{\partial N\overline{R^p}}{\partial t} = \sum_{i=1}^{I} N_i \overline{R_i^p} \left(\frac{\partial \ln N_i}{\partial t} + p \frac{\partial \mu_i}{\partial t} + \frac{p^2}{2} \frac{\partial \sigma_i^2}{\partial t} \right) \qquad (6)$$

Equation (6) can be expressed as a system of equations

$$AX = F \qquad (7)$$

being **X** a vector representing the tendencies of the distribution parameters

135

$$X^T = \left[\frac{\partial \ln N_1}{\partial t}, \frac{\partial \ln N_2}{\partial t}, \dots, \frac{\partial \ln N_I}{\partial t}, \frac{\partial \mu_1}{\partial t}, \frac{\partial \mu_2}{\partial t}, \dots, \frac{\partial \mu_I}{\partial t}, \frac{\partial \sigma^2_1}{\partial t}, \frac{\partial \sigma^2_2}{\partial t}, \dots, \frac{\partial \sigma^2_I}{\partial t} \right] \qquad (8)$$

The coefficient's matrix A is a squared matrix of order v ($v = 3 \times I$) defined as

$$A = \begin{cases} a_{i,j} = N_j \overline{R_j^{i-1}}/(N\overline{R^{i-1}}) \\ a_{i,j+I} = (i-1)N_j \overline{R_j^{i-1}}/(N\overline{R^{i-1}}) \\ a_{i,j+2I} = \frac{1}{2}(i-1)^2 N_j \overline{R_j^{i-1}}/(N\overline{R^{i-1}}) \end{cases} \qquad (9)$$

with $i = 1,2,\dots,v$ and $j = 1,2,\dots,I$. The components of the independent vector F are the tendencies of the total moments of the distributions:

140

$$F^T = \left[\frac{\partial \ln N\overline{R^0}}{\partial t}, \frac{\partial \ln N\overline{R^1}}{\partial t}, \dots, \frac{\partial \ln N\overline{R^{v-1}}}{\partial t} \right] \qquad (10)$$

Both **A** and **F** are normalized in order to achieve a better numerical stability in the solution of the system of equations. The evolution of the values of the parameters of the distribution functions is done by applying a simple forward finite differences scheme (Clark and Hall, 1983)

$$N_i^{k+1} = N_i^k e^{\frac{\partial \ln N_i^k}{\partial t}\Delta t} \qquad (11a)$$

145

$$\mu_i^{k+1} = \mu_i^k + \frac{\partial \mu_i^k}{\partial t}\Delta t \qquad (11b)$$

$$(\sigma^2)_i^{k+1} = (\sigma^2)_i^k + \frac{\partial(\sigma^2)_i^k}{\partial t}\Delta t \quad (11c)$$

2.1 Description of the calculation of the total moment tendencies due to collision-coalescence

The KCE determines the evolution of a DSD due to collision-coalescence. This equation can be expressed in a continuous form dependent of the mass as follows (Pruppacher and Klett, 2010)

150

$$\frac{\partial f}{\partial t} = \int_0^{m/2} f(m - m')f(m')K(m - m'|m')dm' - \int_0^{\infty} f(m)f(m')K(m|m')dm' \quad (12)$$

being $K(m|m')$ the collection kernel. Reformulating eq. (12) in the form of (Thompson, 1968) and in function of radius, we can calculate the total moment tendencies as follows

$$\frac{dN\overline{R^p}}{dt} = \frac{1}{2}\int_0^{\infty}\int_0^{\infty} F^p(r_1, r_2)K(r_1|r_2)f\langle r_1\rangle f\langle r_2\rangle dr_1 dr_2 \quad (13)$$

where

155

$$F^p(r_1, r_2) = \left(r_1^3 + r_2^3\right)^{p/3} - r_1^p - r_2^p \quad (14)$$

$$K\langle r_1|r_2\rangle = \pi(r_1 + r_2)^2 E(r_1, r_2)|V_T(r_1) - V_T(r_2)| \quad (15)$$

Equation (15) represents the hydrodynamic kernel and $E(r_1, r_2)$ stands for the collection efficiencies taken from (Hall, 1980), which is based on a lookup table representing the effectiveness of drop collisions under given environmental conditions. A set of two lognormal distributions (eq. (2)) is used as members of the set at eq. (1). Hence, the prognostic variables under the

160 parameterization formulation will be the corresponding parameters of both distribution functions: N_1, μ_1, σ_1, N_2, μ_2 and σ_2. At this point in the parameterization the total moment tendencies should be calculated either by solving eq. (13) at each time step for all the moments involved, or by searching in a lookup table calculated a priori. Instead, the following section explains in detail the ML approach proposed and its implementation.

3 Machine Learning architecture and training data set

165 Machine Learning methodology can be classified into three main categories, according to the problem at hand: supervised, unsupervised and reinforced learning. In our case, supervised learning is used. Supervised learning algorithms build a mathematical model of a set of data that contains both the inputs and the desired outputs (Russell and Norvig, 2010). Under this classification, there is previous knowledge of the set of input values $\vec{x}_k$, and their corresponding outputs $\vec{y}_k$, , with $k = 1, 2, \ldots, n$, where n is the amount of input values. The objective is to obtain a function $f(\vec{x})$, by means of which the new data

170 $\vec{x}_{new}$ simulates well the output values. The set $\{\vec{x}_k, \vec{y}_k\}$; $k = 1, 2, \ldots, n$ is called the training data set. To test the performance of $f(\vec{x})$, the input and output data is separated into two different data sets : training and validation. As NN are able to fit any

non-linear function (Schmidhuber, 2015), a ML parameterization should approximate reasonable well the solution of the KCE in the form of eq. (13), given enough layers and neurons in the architecture of the network. .

3.1 Neural network architecture

175   Deep Neural Networks are based on artificial neurons. Each neuron receives a set of input data, process it and pass it to an activation function $\sigma(z)$, which returns the activated output (Fig. 1). The activation value of neuron i in layer l is denoted by a_i^l and is determined as

$$a_i^l = \sigma(z_i^l) \qquad (16)$$

$$z_i^l = b_i^l + \sum_{j=!}^{m_{l-1}} W_{i,j}^l a_i^{l-1} \qquad (17)$$

180   In eq. (17), b_i^l is the bias, $W_{i,j}^l$ is the ponderation weight, m_{l-1} the number of neurons in layer l-1, and $\sigma(z)$ is the activation function. Hence, a NN could be defined as a set of input values $(\vec{x})$, bias values $(\vec{b})$ and weights $(\vec{W})$ integrated in a functional form, i.e. $\vec{y}(\vec{x}, \vec{W}, \vec{b})$, and its training procedure consists on minimizing an error function (known as loss function), by optimizing the weights and biases for the available training data. A commonly used loss function is the regression mean squared error (MSE). Hence, we need a minimization algorithm to process the following expression

185    $$C(\vec{W}, \vec{b}) = \frac{1}{2n} \sum_k \left\| \vec{y}(\vec{x}_k, \vec{W}, \vec{b}) - \vec{y}_k \right\|^2 \qquad (18)$$
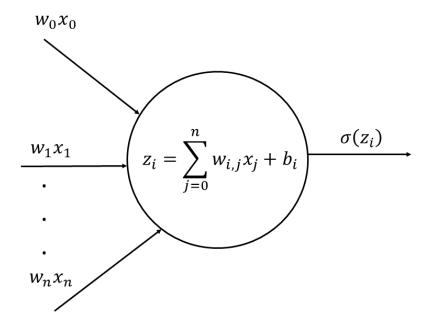

Figure 1: Schematic representation of an artificial neuron.

The selected algorithm for minimization of the loss function (eq. (18)) is the bayesian regularization, which updates the weights and biases values according to Levenberg-Marquardt optimization. Backpropagation is used to calculate the Jacobian of the

190 performance with respect to the weight and bias variables (Dan Foresee and Hagan, 1997; MacKay, 1992).

The used DNN was developed using MATLAB. It is conformed by one layer which receives the input data (input layer), three intermediate layers (hidden layers) with 20 neurons each and an output layer with a single neuron which returns the simulated target values (Fig. 2).
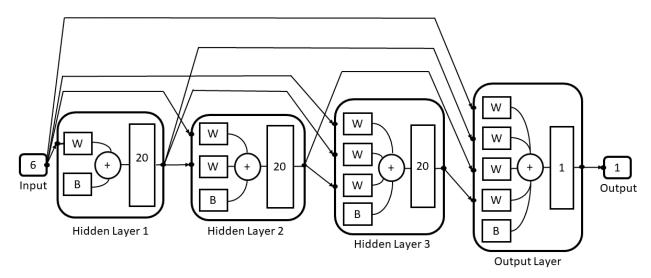

195 **Figure 2: Schematic representation of the architecture of the trained neural network used to calculate the total moment tendencies. The neural network receives six inputs, then processes them by means of three hidden layers of 20 neurons each, and an output layer with a single neuron and one output value.**

3.2 Generation of the training and validation data sets

The training procedure consists on feeding the DNN with six input values corresponding to the distribution parameters of each

200 distribution and the total moment tendency for the p-th order obtained from eq. (13) as target. The NN training algorithm then processes those values in order to establish the relationships between the data provided. This process is repeated until all input and target data is processed. The resulting trained DNN should be able to estimate the total moment tendencies from a given set of distribution parameters that falls within the ranges of the training variables. A schematic representation of the trained NN with the inputs and output is shown in Fig. 3.
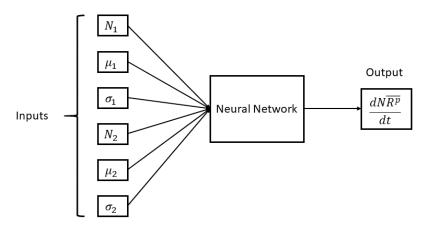
Geoscientific
Model Development
Discussions

Figure 3: Neural network parameterization inputs and output. The input data are the six distribution parameters $(N_1, \mu_1, \sigma_1, N_2, \mu_2$ and $\sigma_2)$ needed to feed eq. (13), while the output is the *p*-th order total moment tendency $\left(\frac{dN\overline{R^p}}{dt}\right)$.

In order to generate the training and validation data sets, 100000 drop spectra derived from the input variables are employed, over a wide range of distribution parameters ($N_1, \mu_1, \sigma_1, N_2, \mu_2$ and σ_2) (See Fig. 4). Those input parameters will be used to calculate the total moment rates from eq. (13) and train the DNN. Five DNNs will be trained, one for each total moment tendency involved in the formulation of the parameterization (moment orders ranged from 0 to 5). The same training input parameters are used to train all NNs, varying only the target values corresponding to the total moment tendencies of each order. The physical variables related to the input parameters are shown in Fig. 5 for a better representation of the generated training clouds. The training and validation data was created using an uniformly distributed random number generator, with means and standard deviations shown in Table 1, as well as the ranges of each predictor variable.

Table 1: Statistical description of the input values used in the training and validation data sets. The means, standard deviation and ranges are shown for each input variable.

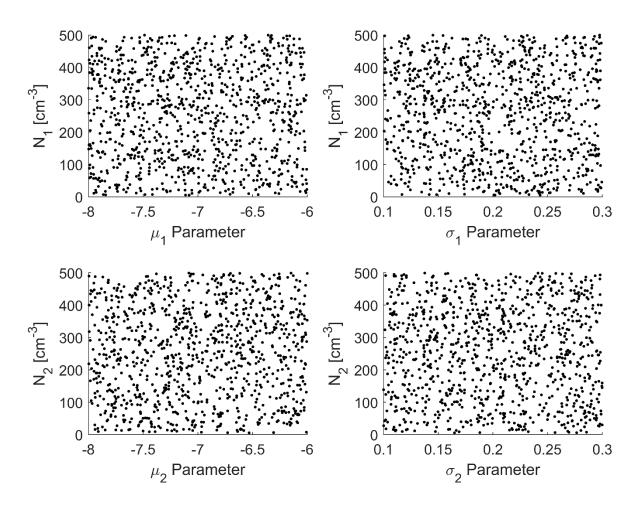| Input Variable	Mean	Standard Deviation	Range
Concentration (N)	250.80	144.13	[1.00; 500.00]
μ Parameter	-7.00	0.58	[-8.00; -6.00]
σ Parameter	0.20	0.06	[0.10; 0.30]

220

Figure 4: Scatterplot of the input parameters ($N_1, \mu_1, \sigma_1, N_2, \mu_2$ and σ_2) of the lognormal distributions used to calculate the total moment tendencies. Parameters for distribution function f_1 are shown in the first row, while parameters for f_2 are shown in the second row. Only every 100th data point is shown.

Figure 5 shows that within the ranges of the training data (concentration from 1 cm^{-3} to 500 cm^{-3}), the corresponding liquid

225　water contents (LWC) are between 10^{-10} g cm^{-3} and 10^{-4} g cm^{-3}, with the majority of the data concentrated between the limits of 10^{-8} g cm^{-3} and 10^{-5} g cm^{-3}. Those values cover a sufficiently wide range of liquid water content to adequately represent warm clouds within the parameterization.
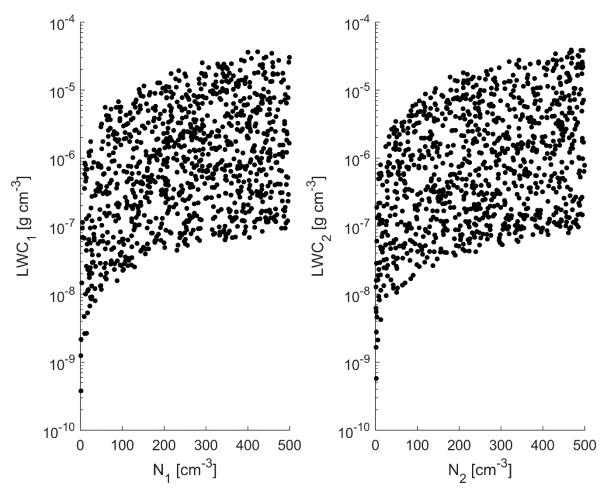
10

Figure 5: Scatterplot of liquid water content (LWC) calculated from the input parameters of f_1 (left) and f_2 (right) vs drop number concentration. The LWC values are obtained from the statistical moment of order 3 using the parameters depicted in Fig. (4), and were calculated from eq. (4). Only every 100th data point is shown.

3.3 Training and validation of the Deep Neural Network

From the available data, 80 % was employed in training the DNN, and the remaining 20 % was used for validation purposes. The total moment tendencies (eq. 13) were solved using a trapezoidal rule, over a logarithmic radius grid between the ranges of $1\,\mu m \leq r \leq 10^4\,\mu m$. The results of solving eq. (13) are called the target values. The means and standard deviation for each calculated total moment rate are shown in Table 2.

Table 2: Total moment tendencies (target values) means and standard deviations for each statistical moment used. The data is calculated from eq. (13) with the distribution parameters ($N_1, \mu_1, \sigma_1, N_2, \mu_2$ and σ_2) as input values.

Total Moment Order	Mean	Standard Deviation

M0	-0.0021	0.0014
M1	-0.0015	0.0011
M2	-0.0009	0.0006
M3	0.0000	0.0000
M4	0.0011	0.0007
M5	0.0024	0.0016

240 Both input and target values are normalized as follows

$$x_{norm} = \frac{x - \bar{x}}{\sigma} \qquad (19)$$

The input and target values require a normalization process to facilitate the work of the optimization algorithm. All the nodes in each layer of the DNN use the MSE as loss function. The training procedure for a NN consist of processing a fragment of the total training data through the network learning architecture , then determining the prognostic error and the gradient of the

245 loss function (MSE) back through the network in order to update the weight values. This algorithm is repeated via an iterative process over all training data until the performance index (MSE) is small enough or a predefined number of passes through all data are completed. One pass through all training data is known as an epoch. In this case, a maximum number of 1000 epochs is established, and a minimum value of 10^{-7} is considered for the gradient function.

Five DNN were trained, one for each total moment tendency involved in the formulation of the parameterization (moment

250 orders ranged from 0 to 5). A variant of the training process, known as cascade-forward neural network training, was employed. The main difference with the standard training procedure is the fact that includes a connection from the input and every previous layer to following layers (see Fig. 2). As with feed-forward networks, a two-or more layer cascade network can learn any finite input-target relationship arbitrarily well, given enough hidden neurons. The total moment tendencies for the statistical moment of order 3 was not calculated because the collision-coalescence process does not affect total mass.

255 Performance (MSE) training records for the total moment tendencies calculated from eq. (13) are depicted in Fig. 6. The speed of convergence is similar in all cases, and all networks converged at epoch 1000. This occurs due to the gradient value never was below the minimum, so the training process kept refining the results until it reached the maximum number of epochs previously defined. Despite that, a good performance was achieved, being the MSE in the order of 10^{-4} for all orders of the total moment tendencies, as shown in Table 3, where the best MSE values for each trained DNN are manifested in detail.

260 **Table 3: Best performance in the training process of the DNNs. The shown data corresponds to the total moment tendencies obtained from the trained neural networks, with input values and reference targets taken from the validation data set. The performance measure is the Mean Squared Error (MSE).**

Total Moment Order	Best Performance (MSE)
M0	2.59e-04
M1	3.49e-04

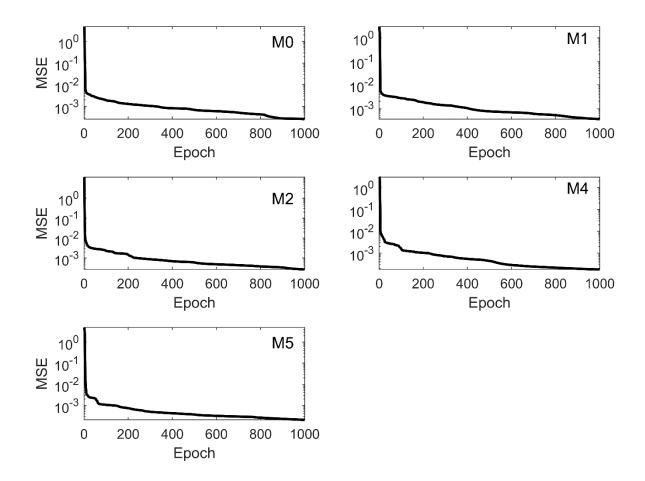| M2 | 2.68e-04 |
| M4 | 1.80e-04 |
| M5 | 2.05e-04 |

Figure 6: Performance training records of total moment tendencies for the moments from order 0 to 5. The shown data corresponds to the total moment tendencies obtained from the trained neural networks, with input values and reference targets taken from the validation data set. The performance measure is the Mean Square Error (MSE). The performance for the moment of order 3 is not included because the collision-coalescence process does not affect total mass.

Regression plots for the trained networks are depicted in Fig. 7. The comparison is established between the outputs obtained from evaluating the trained neural networks using the validation inputs and the targets from the validation data set corresponding to each of the total moment tendencies obtained from eq. (13). Minor differences can be appreciated from the graphics, with the trained DNN models overestimating or underestimating the actual values. However, a good agreement was

reached for all trained DNN, with the predicted values from the DNN matching the actual output from the solution of eq. (13)

275 with a coefficient of correlation between 0.9998 and 0.9999 in all cases. The axis ranges of the graphics varies because the

plotted data is not normalized, thus, there are different ranges for each of the calculated total moment tendencies.
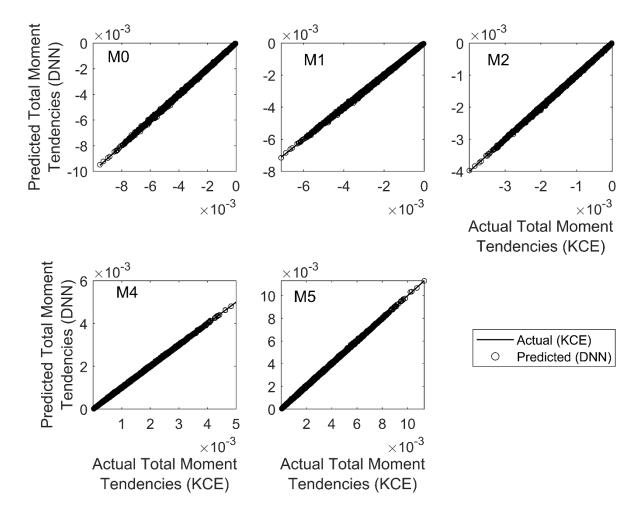

Figure 7: Regression plots for the five DNN trained. The comparison is established between the outputs obtained from evaluating the trained neural networks using the validation inputs and the targets from the validation data set
280 **corresponding to each of the total moment tendencies obtained from eq. (13). The order of the statistical moments range from 0 to 5. Regression for the moment of order 3 is not shown because the collision-coalescence process does not affect total mass. The *y* axis varies for each subplot because the plotted data is not normalized.**

Experiments with not normalized training data were performed, yielding results with MSE at least an order of magnitude

higher. Those results are not shown in the present article due to the lower accuracy of the regression outputs. The codes for

285 generating the training and validation data sets, and training of the neural networks themselves can be found at (Rodríguez-

Genó and Alfonso, 2021c)

4 Description of the reference model

To obtain a reference solution, the explicit model developed by (Bott, 1998a) was employed. This scheme is conservative with respect to mass and very efficient computationally speaking. It is based on the numerical integration of the KCE which, rewritten in a more convenient way, is expressed as shown below:

$$\frac{dn(x,t)}{dt} = \int_{x_0}^{x_1} n(x_c,t)K(x_c,x')n(x',t)dx' - \int_{x_0}^{\infty} n(x,t)K(x,x')n(x',t)dx' \tag{20}$$

where $n(x,t)$ stands for the DSD at time t and $K(x_c,x')$ represents the collection kernel. In order to simplify the calculations, the mass density function $g(y,t)$ is used (Berry, 1967):

$$g(y,t)dy = xn(x,t)dx \tag{21}$$

$$n(x,t) = \frac{1}{3x^2}g(y,t) \tag{22}$$

where $y = \ln r$ and r is the radius of a drop of mass x. By substituting (21) in (20) we obtain the KCE for the mass density function (Bott, 1998a)

$$\frac{dg(y,t)}{dt} = \int_{y_0}^{y_1} \frac{x^2}{x_c^2 x'}g(y_c,t)K(y_c,y')g(y',t)dy' - \int_{y_0}^{\infty} g(y,t)\frac{K(y,y')}{x'}g(y',t)dy' \tag{23}$$

The first integral of the right-hand side of eq. (23) represents the gain of drops of mass x due to collision-coalescence of two smaller droplets, while the second integral portray the loss of drops of mass x being captured by bigger drops (Bott, 1998a). For the numerical solution of eq. (23), a logarithmic equidistant mass grid is used, and is generated as

$$x_{k+1} = \alpha x_k, \qquad k = 1,2,\dots,m \tag{24}$$

where m is the total number of grid points. The original code for this explicit model can be found at (Bott, 1998b), and can be used with authorization of the author.

5 Initial conditions and experiment design

A total time of $t = 900\ s$ (15 minutes) is simulated for both the parameterized and reference models, with a time step of $\Delta t = 0.1\ s$. The initial parameters for the distribution functions of the parameterized model were established as shown in Table 4, following (Clark, 1976):

Table 4: Initial parameters for the distribution functions of the parameterized model. Each distribution is characterized by a concentration parameter (N), expected value (μ) and standard deviation (σ). The initial parameters are shown for the two lognormal distribution functions employed in the formulation of the parameterization.

Parameter	f_1	f_2
N	190 cm^{-3}	10 cm^{-3}
μ	-7.1505	-6.5219

σ	0.1980	0.1980

The initial spectrum for the reference solution was calculated from these parameters to ensure the same initial conditions for both models. A 300 points logarithmic equidistant grid was generated for the integration of the reference model, with radii values in the range of $0.25\ \mu m \leq r \leq 2.6 \times 10^4\ \mu m$. Equations (21) and (22) were used to transform the output of both models

315 to make them compatible for comparison, while the bulk quantities from the reference models were integrated from the calculated spectra.

6 Discussion of results

The results shown in this section were obtained using the parameterized model COLNETv1.0.0. The source code can be found at (Rodríguez-Genó and Alfonso, 2021a), while the scripts for reproducing the figures are archived at (Rodríguez-Genó and

320 Alfonso, 2021b)

6.1 Spectra comparison

The output of this parameterized model are the updated distribution parameters at every time step (N_1, μ_1 , σ_1, N_2, μ_2 and σ_2). The physical variables related to the moments of the distributions, such as mean radius or liquid water content (LWC) are diagnosed from those parameters. Besides, we can derive the shape and scale of the drop spectrum at any given time, by

325 integrating the functions and supported on the evolution of its parameters.

Figure 8 shows a comparison between the mass density spectra derived from the parameterized and explicit models for three chosen times (300 s, 600 s and 900 s).
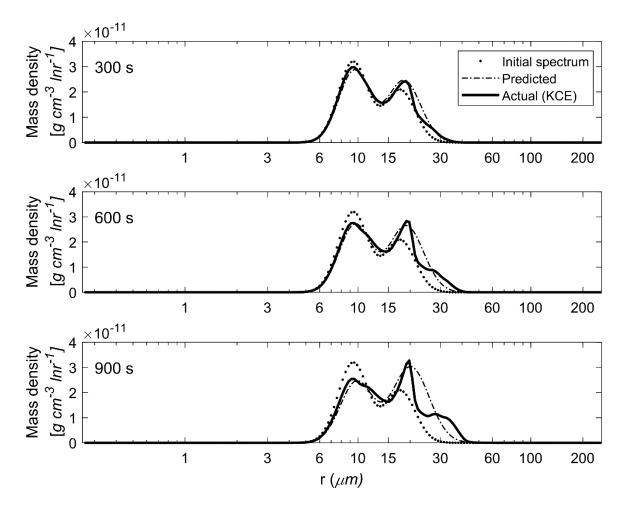
Figure 8: Mass density spectra from the parameterized model (Predicted) and the reference solution (Actual). The
330 **represented times are 300 s, 600 s and 900 s, from top to bottom. Equation (21) was used to transform the drop number**
concentration spectra from the parameterized model to the mass density spectra.

At 300 s (first row of Fig. 8), there is a slow development of the total spectrum, with a clear mass transfer from distribution
function f_1 to f_2. The parameter-generated spectrum fits well the reference solution, with a slight overestimation of the
maximum mass in the second mode. The mean radius of the distributions are well represented by the parameterization. At 600
335 s and 900 s (second and third row of Fig. 8), there is a development of a third mode in the evolution of the KCE-generated
spectra, that is reproduced by the parameterization as a wider second mode, representing well the mean radius and mass
distribution. The first mode is accurately represented at those times. An increase on mean radius can be observed, due to the
effect of the collision-coalescence process.

Figure 9 shows a comparison between the drop number concentration spectra derived from the parameterized and explicit
340 models for three chosen times (300 s, 600 s and 900 s). A generally good agreement is appreciated at all times, with the

parameterization-generated spectra slightly underestimating the results from the explicit model. As the collision-coalescence process decreases the drop number concentration, there is not a noticeable increase in the number of drops in the second mode of the distributions. However, an increase in the mean radius is observed, that is consistent with the behaviour described in Fig. 8, where is a related mass transfer between both distribution functions.
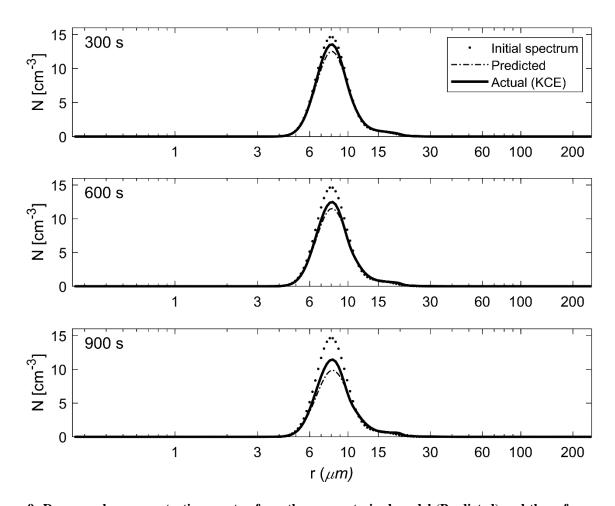

345

Figure 9: Drop number concentration spectra from the parameterized model (Predicted) and the reference solution (Actual). The selected times are 300 s, 600 s and 900 s, from top to bottom. Equation (22) was used to transform the mass density spectra from the reference solution to the drop number concentration spectra.

6.2 Bulk quantities comparison

350 Figure 10 shows a comparison of two main bulk quantities (total number concentration and mean radius) obtained from the parameterization and the reference model. The concentration and mean radius of the reference solution were obtained by integrating the drop number concentration spectra for the corresponding moment order (0 and 1 respectively). As expected, number concentration decreases with time, due to the coalescence of drops, ranging from an initial value of 200 drops by cm⁻

18

[3] to around 160 by cm[-3] in the KCE model. The predicted concentration underestimates the actual values in most of the

355 simulation, with the differences reaching 10 drops by cm[-3] at 900 s. A similar behaviour is observed in the mean radius results,

with a growth in the drop size consistent with the decreasing values on the drop number, but differences are small with the

highest error with a value of $5 \times 10^{-5}\ cm$. However, the results fit well enough the reference solution to consider including

this microphysical parameterization in a weather model. This consistent behaviour of the mean radius and number

concentration values points to a conservation of mass, a compulsory condition in the collision-coalescence process.
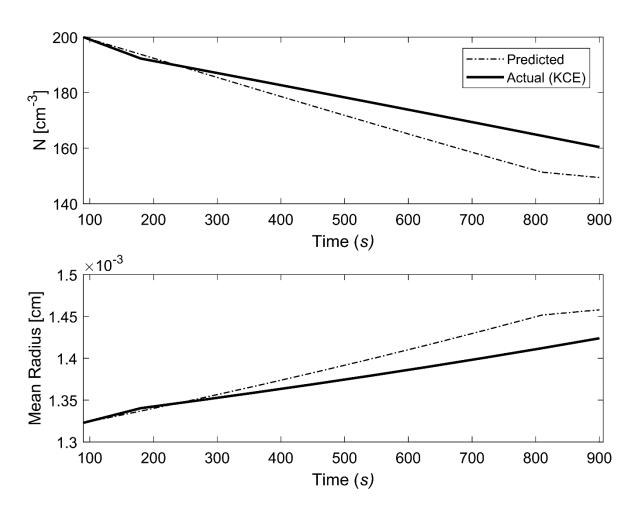

Figure 10: Drop number concentration (top) and mean radius (bottom) comparison with the reference solution. The concentration and mean radius of the reference solution were obtained by integrating the drop number concentration spectra for the corresponding moment order (0 and 1 respectively). The data points are plotted every 60 s.

Figure 11 depicts the evolution of two main bulk quantities (drop number concentration and liquid water content) for the

365 individual distributions that conform the parameterization (f_1 and f_2), as well as the combined (total) values of the variables

(calculated as $f_1 + f_2$). Regarding concentration, a decrease in f_1 values is observed, due to the coalescence process, while a

consistent increase in f_2 is also appreciated. The increase in the concentration of f_2 is not as marked due to the collision-coalescence process as well. However, a general decrease of the total concentration value (the only distribution parameter with physical meaning in the formulation of the parameterization) represents well the theory and observations of the parameterized process.

The liquid water content LWC) values (diagnosed) are depicted to verify that mass is conserved under the formulation of the parameterization. The liquid water content of each of the distribution functions (f_1 and f_2) were obtained from the corresponding moment (order 3) calculated from eq. (4). Effectively, the mass content retains a value of $6.1739 \times 10^{-7}\ g\ cm^{-3}$ during the entire simulation, with a proportional mass transfer between f_1 and f_2.
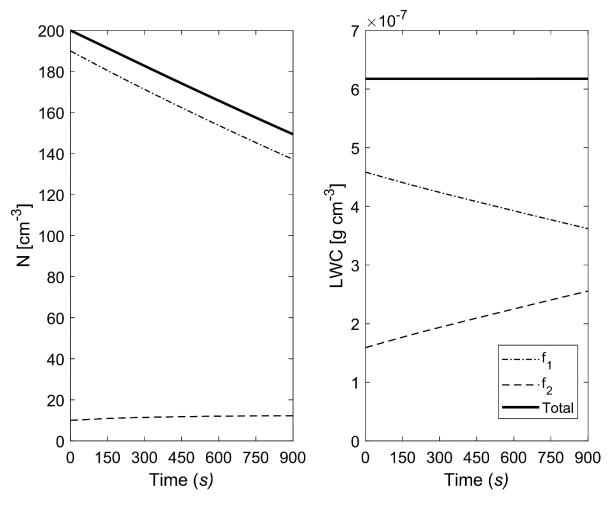

Figure 11: Evolution of drop number concentration N (left) and liquid water content LWC (right) of the individual distributions that conform the parameterized model. The liquid water content of each of the distribution functions (f_1 and f_2) were obtained from the corresponding moment (order 3) calculated from eq. (4). The combined (total) values of the variables are also shown and were calculated from eq. (5).

Geoscientific
Model Development
Discussions

EGU

6.3 Total moment errors

An analysis of the predicted total moments was performed with the objective of further test the precision of the DNN collision-coalescence parameterization, due to the importance of the statistical moments in calculating physical variables such as mean radius and LWC. Table 5 shows the mean percent errors of the calculations of the total moments of the parameterization. The percent error is taken relative to the moments of the reference solution. The data was obtained by calculating the mean of the percent errors of the entire simulation. The moments of the reference solution were calculated by integrating the reference drop number concentration spectra using eq. (3), while the total moments from the parameterized model were calculated using the predicted distribution parameters and solving eq. (5). A reasonable degree of accuracy was achieved, with the mean error never surpassing the 4 %. However, the data shows that the total moments of order 0 to 2 are usually underestimated, while those of order 4 and 5 are slightly overestimated. This could result in the calculations of drop number concentration values lower than the actual ones, as seen in Fig. 10.

Table 5: Total moment mean errors. The percent error is taken relative to the moments of the reference solution. The shown data was obtained by calculating the mean of the percent errors of the entire simulation. The moments of the reference solution were calculated by integrating the reference drop number concentration spectra using eq. (3), while the total moments from the parameterized model were calculated using the predicted distribution parameters to solve eq. (5).

Total Moment Order	Mean Percent Error
M0	-3.3479
M1	-2.6437
M2	-1.4969
M3	0
M4	1.1249
M5	0.7205

Figure 12 shows the time evolution of the percent error of the total moments throughout the parameterization simulation. The percent error is taken relative to the moments of the reference solution. The moments of the reference solution were calculated by integrating the reference drop number concentration spectra using eq. (3), while the total moments from the parameterized model were calculated using the predicted distribution parameters to solve eq. (5). The error of total moment of order 3 is zero during the entire simulation because mass is not affected by the collision-coalescence process. The total moments from order 0, 1 and 2 overestimate the reference solution in the first 300 s of simulation, underestimating them for the rest of the parameterization run, with the percent error reaching a minimum value of -8 %. The opposite behaviour is appreciated for the total moments of order 4 and 5, where they initially underestimate the reference solution, overestimating it for the rest of the simulation. However, for these orders the percent error is usually lower, with a maximum of 4 %. Generally, the

parameterization is performing well, with the percent error never reaching the 10 % threshold. However, further analysis on this topic is recommended, to improve the accuracy of the parameterization.
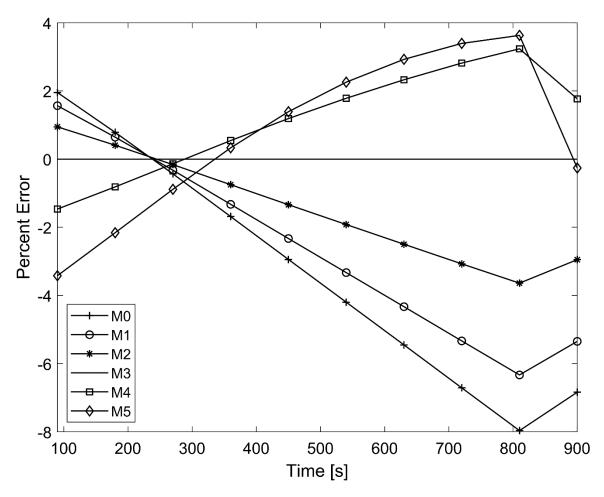

Figure 12: Time evolution of the errors corresponding to the predicted moments from the parameterized model. The percent error is taken relative to the moments of the reference solution. The moments of the reference solution were calculated by integrating the reference drop number concentration spectra using eq. (3), while the total moments from the parameterized model were calculated using the predicted distribution parameters to solve eq. (5).

7 Conclusions

A hybrid parameterization for the process of collision-coalescence based on the methodology of basis functions employing a linear combination of two lognormal distributions was developed and implemented, with all the parameters of the distributions derived from the total moment tendencies calculated by means of five trained deep neural networks. By doing this, we obtained a parameterized model that determines the distribution parameters evolution, hence, the evolution of the DSD. The physical

22

variables related to the moments of the distributions are diagnosed. Within the framework of this parameterized model, there is no artificial separation of the DSD, thus, terms such as cloud droplet and raindrops loose meaning as there are just "drops".

420 Instead, we consider a full set of distribution parameters for each of the distribution functions that are considered in the formulation of the parameterization, in order to describe the DSD in radius space. This kind of microphysical parameterization allows the use of an arbitrary number of statistical density functions in linear combination to reproduce the drop spectrum.

One experiment was performed to test the proposed ML formulation at the initial stages of cloud formation. The simulation results showed good agreement when compared to a reference solution, for both the predicted DSD and the bulk quantities

425 considered. The total moment tendencies were well predicted using the trained DNNs, improving the computational performance of the original formulation. An analysis of the accuracy of the predicted total moments of the parameterized model was performed, with the percent error relative to the moments of the reference solution never reaching the 10 % threshold. However, there is room for improvement in the calculations of the total moments, being the recommendation of the authors to retrain the DNNs with a finer parameter domain, and with a larger range of values in order to cover all possible

430 combination of parameters. In addition, the use of Machine Learning eliminated the requirement of integrating the total moment tendencies at each time step, and the use of memory expensive lookup tables for each predicted moment is no longer needed under this formulation, fact that leads to an improvement in precision, compared with standard interpolation methods. The presented way to simulate the evolution of the droplet spectra due to collision-coalescence falls within the framework of the methodology of series of basis functions developed by (Clark, 1976; Clark and Hall, 1983). Under this modelling

435 philosophy, a dynamic framework has been established in (Rodríguez-Genó and Alfonso, 2021d). To obtain a full warm cloud model, an extension of this neural network algorithm applied to condensation is proposed, following the same methodology of series of basis functions. A parameterization scheme such as this could be included in regional weather and climate models, as its initial conditions can be calculated from the ones needed by for more traditional bulk models.

Author contributions

440 Léster Alfonso performed the conceptualization of the article, methodology, funding acquisition, resources, supervision and reviewing and editing the original draft. Camilo Fernando Rodríguez Genó realized the formal analysis, investigation, software and code development, validation, visualization, and writing the original draft preparation.

Code availability

The current version of *COLNET (COLNETv1.0.0)* used to produce the results presented in this paper is archived on Zenodo

445 (https://doi.org/10.5281/zenodo.4740061) under license *GNU Affero General Public License v3.0 or later*, as well as all dependency scripts to run the model. The outputs of the model used to generate the figures included in the present paper are also included. The scripts used in the generation of training data sets and for training the neural networks used in

Geoscientific
Model Development
Discussions

COLNETv1.0.0 can be found on Zenodo (https://doi.org/10.5281/zenodo.4740129), while the codes for plotting the figures are stored at https://doi.org/10.5281/zenodo.4740184. The original code of the explicit bin model is archived at
450   https://www2.meteo.uni-bonn.de/forschung/gruppen/tgwww/people/abott/fortran/coad1d.f, and have been used with the permission of the author. The model and related scripts were developed using MATLAB R2020a under license number 40816183, with exception of the explicit model, which is coded on FORTRAN 77.

Competing interests

455   The authors declare that they have no conflict of interests.

465   **References**

Alfonso, L. and Zamora, J. M.: A two-moment machine learning parameterization of the autoconversion process, Atmos. Res., 249, 105269, doi:10.1016/j.atmosres.2020.105269, 2021.

Alfonso, L., Raga, G. B. and Baumgardner, D.: The validity of the kinetic collection equation revisited, Atmos. Chem. Phys., 8(4), 969–982, doi:10.5194/acp-8-969-2008, 2008.

470   Alfonso, L., Raga, G. and Baumgardner, D.: A Monte Carlo framework to simulate multicomponent droplet growth by stochastic coalescence, in Applications of Monte Carlo method in science and engineering, edited by S. Mordechai, InTech., doi: 10.5772/14522, 2011.

Berry, E. X.: Cloud droplet growth by collection, J. Atmos. Sci., 24(6), 688–701, doi:10.1175/1520-0469(1967)024<0688:CDGBC>2.0.CO;2, 1967.

475   Berry, E. X. and Reinhardt, R. L.: An analysis of cloud drop growth by collection: Part I. Double distributions, J. Atmos. Sci., 31(7), 1814–1824, doi:10.1175/1520-0469(1974)031<1814:AAOCDG>2.0.CO;2, 1974.

Bott, A.: A flux method for the numerical solution of the stochastic collection equation, J. Atmos. Sci., 55(13), 2284–2293,

doi:10.1175/1520-0469(1998)055<2284:AFMFTN>2.0.CO;2, 1998a.

Bott, A.: Program for the solution of the stochastic coalescence equation: One-dimensional cloud microphysics., [online]
480   Available from: https://www2.meteo.uni-bonn.de/forschung/gruppen/tgwww/people/abott/fortran/coad1d.f, 1998b.

Clark, T. L.: Use of log-normal distributions for numerical calculations of condensation and collection, J. Atmos. Sci., 33(5),
810–821, doi:10.1175/1520-0469(1976)033<0810:UOLNDF>2.0.CO;2, 1976.

Clark, T. L. and Hall, W. D.: A cloud physical parameterization method using movable basis functions: Stochastic coalescence
parcel calculations, J. Atmos. Sci., 40(7), 1709–1728, doi:10.1175/1520-0469(1983)040<1709:ACPPMU>2.0.CO;2, 1983.

485   Cohard, J.-M. and Pinty, J.-P.: A comprehensive two-moment warm microphysical bulk scheme. I: Description and tests, Q.
J. R. Meteorol. Soc., 126(566), 1815–1842, doi:10.1256/smsqj.56613, 2000.

Dan Foresee, F. and Hagan, M. T.: Gauss-Newton approximation to Bayesian learning, in Proceedings of International
Conference on Neural Networks (ICNN'97), vol. 3, pp. 1930–1935, IEEE, Houston., doi: 10.1109/ICNN.1997.614194, 1997.

Dekkers, P. J. and Friedlander, S. K.: The self-preserving size distribution theory: I. Effects of the Knudsen number on aerosol
490   agglomerate growth, J. Colloid Interface Sci., 248(2), 295–305, doi:10.1006/jcis.2002.8212, 2002.

Feingold, G. and Levin, Z.: The lognormal fit to raindrop spectra from frontal convective clouds in Israel, J. Clim. Appl.
Meteorol., 25(10), 1346–1363, doi:10.1175/1520-0450(1986)025<1346:TLFTRS>2.0.CO;2, 1986.

Feingold, G., Walko, R. L., Stevens, B. and Cotton, W. R.: Simulations of marine stratocumulus using a new microphysical
parameterization scheme, Atmos. Res., 47–48, 505–528, doi:10.1016/S0169-8095(98)00058-1, 1998.

495   Gillespie, D. T.: The stochastic coalescence model for cloud droplet growth, J. Atmos. Sci., 29(8), 1496–1510,
doi:10.1175/1520-0469(1972)029<1496:TSCMFC>2.0.CO;2, 1972.

Hall, W. D.: A detailed microphysical model within a two-dimensional dynamic framework: Model description and
preliminary results, J. Atmos. Sci., 37(11), 2486–2507, doi:10.1175/1520-0469(1980)037<2486:ADMMWA>2.0.CO;2, 1980.

Huang, W. T. K.: Reformulation of the Milbrandt and Yau microphysics parameterization: A triple-moment lognormal scheme,
500   McGill University, Montréal, Québec., url: https://escholarship.mcgill.ca/concern/theses/3f4628385, 2014.

Kessler, E.: On the distribution and continuity of water substance in atmospheric circulations, in Meteorological Monographs,
vol. 10, pp. 1–84, American Meteorological Society, Boston, MA., doi: 10.1007/978-1-935704-36-2_1, 1969.

Khain, A., Pokrovsky, A., Pinsky, M., Seifert, A. and Phillips, V.: Simulation of effects of atmospheric aerosols on deep
turbulent convective clouds using a spectral microphysics mixed-phase cumulus cloud model. Part I: Model description and
505   possible applications, J. Atmos. Sci., 61(24), 2963–2982, doi:10.1175/JAS-3350.1, 2004.

Khain, A. P., Lynn, B. and Dudhia, J.: Aerosol effects on intensity of landfalling hurricanes as seen from simulations with the
WRF model with spectral bin microphysics, J. Atmos. Sci., 67(2), 365–384, doi:10.1175/2009JAS3210.1, 2010.

Khairoutdinov, M. and Kogan, Y.: A new cloud physics parameterization in a large-eddy simulation model of marine
stratocumulus, Mon. Weather Rev., 128(1), 229–243, doi:10.1175/1520-0493(2000)128<0229:ANCPPI>2.0.CO;2, 2000.

510   Koza, J. R., Bennett, F. H., Andre, D. and Keane, M. A.: Automated design of both the topology and sizing of analog electrical
circuits using genetic programming, in Artificial Intelligence in Design '96, edited by J. S. Gero and F. Sudweeks, pp. 151–
170, Springer Netherlands, Dordrecht., doi: 10.1007/978-94-009-0279-4_9, 1996.

Lee, K. W., Chen, J. and Gieseke, J. A.: Log-normally preserving size distribution for Brownian coagulation in the free-

molecule regime, Aerosol Sci. Technol., 3(1), 53–62, doi:10.1080/02786828408958993, 1984.

515 Lee, K. W., Lee, Y. J. and Han, D. S.: The log-normal size distribution theory for Brownian coagulation in the low Knudsen number regime, J. Colloid Interface Sci., 188(2), 486–492, doi:10.1006/jcis.1997.4773, 1997.

Lim, K.-S. S. and Hong, S.-Y.: Development of an effective double-moment cloud microphysics scheme with prognostic cloud condensation nuclei (CCN) for weather and climate models, Mon. Weather Rev., 138(5), 1587–1612, doi:10.1175/2009MWR2968.1, 2010.

520 Long, A. B.: Solutions to the droplet collection equation for polynomial kernels, J. Atmos. Sci., 31(4), 1040–1052, doi:10.1175/1520-0469(1974)031<1040:STTDCE>2.0.CO;2, 1974.

MacKay, D. J. C.: Bayesian interpolation, Neural Comput., 4(3), 415–447, doi:10.1162/neco.1992.4.3.415, 1992.

Marshall, J. S. and Palmer, W. M. K.: The distribution of raindrops with size, J. Meteorol., 5(4), 165–166, doi:10.1175/1520-0469(1948)005<0165:TDORWS>2.0.CO;2, 1948.

525 Milbrandt, J. A. and McTaggart-Cowan, R.: Sedimentation-induced errors in bulk microphysics schemes, J. Atmos. Sci., 67(12), 3931–3948, doi:10.1175/2010JAS3541.1, 2010.

Milbrandt, J. A. and Yau, M. K.: A multimoment bulk microphysics parameterization. Part I: Analysis of the role of the spectral shape parameter, J. Atmos. Sci., 62(9), 3051–3064, doi:10.1175/JAS3534.1, 2005.

Mitchell, T.: Machine Learning, McGraw Hill, New York. [online] Available from: http://www.cs.cmu.edu/~tom/mlbook.html
530 (Accessed 4 May 2021), isbn: 0070428077, 1997.

Morrison, H., Thompson, G. and Tatarskii, V.: Impact of cloud microphysics on the development of trailing stratiform precipitation in a simulated squall line: Comparison of one- and two-moment schemes, Mon. Weather Rev., 137(3), 991–1007, doi:10.1175/2008MWR2556.1, 2009.

Noh, Y., Oh, D., Hoffmann, F. and Raasch, S.: A cloud microphysics parameterization for shallow cumulus clouds based on
535 lagrangian cloud model simulations, J. Atmos. Sci., 75(11), 4031–4047, doi:10.1175/JAS-D-18-0080.1, 2018.

Pruppacher, H. R. and Klett, J. D.: Microphysics of clouds and precipitation, 2nd ed., Springer Netherlands, Dordrecht., doi: 10.1007/978-0-306-48100-0, 2010.

Rodríguez-Genó, C. F. and Alfonso, L.: COLNETv1.0.0, , doi:10.5281/ZENODO.4740061, 2021a.

Rodríguez-Genó, C. F. and Alfonso, L.: COLNETv1.0.0 graphics generation scripts, , doi:10.5281/ZENODO.4740184, 2021b.

540 Rodríguez-Genó, C. F. and Alfonso, L.: COLNETv1.0.0 neural network training scripts, , doi:10.5281/ZENODO.4740129, 2021c.

Rodríguez-Genó, C. F. and Alfonso, L.: Sedimentation calculations within an Eulerian framework using series of basis functions, Q. J. R. Meteorol. Soc., 147(736), 2053–2066, doi:10.1002/qj.4009, 2021d.

Russell, S. and Norvig, P.: Artificial intelligence: a modern approach, Third Edit., Prentice Hall., isbn: 9780262018258, 2010.

545 Schmidhuber, J.: Deep learning in neural networks: An overview, Neural Networks, 61, 85–117, doi:10.1016/j.neunet.2014.09.003, 2015.

Seifert, A. and Beheng, K. D.: A double-moment parameterization for simulating autoconversion, accretion and selfcollection, Atmos. Res., 59–60, 265–281, doi:10.1016/S0169-8095(01)00126-0, 2001.

Seifert, A. and Rasp, S.: Potential and limitations of machine learning for modeling warm-rain cloud microphysical processes,
550   J. Adv. Model. Earth Syst., 12(12), doi:10.1029/2020MS002301, 2020.

von Smoluchowski, M.: Drei Vorträge über Diffusion, Brownsche Molekularbewegung und Koagulation von Kolloidteilchen
[Part 1], Phys. Zeitschrift, 17(22), 557–571 [online] Available from: https://fbc.pionier.net.pl/id/oai:jbc.bj.uj.edu.pl:387533,
1916a.

von Smoluchowski, M.: Drei Vorträge über Diffusion, Brownsche Molekularbewegung und Koagulation von Kolloidteilchen
555   [Part 2], Phys. Zeitschrift, 17(23), 585–599 [online] Available from: https://fbc.pionier.net.pl/id/oai:jbc.bj.uj.edu.pl:387534,
1916b.

Sobhani, N., Loft, R., Gettelman, A., Chen, C. C. and Gagne, D. J., I.: Using machine learning to emulate critical cloud
microphysical processes, in American Geophysical Union, Fall Meeting 2018, pp. IN13C-0686. [online] Available from:
https://ui.adsabs.harvard.edu/abs/2018AGUFMIN13C0686S (Accessed 29 October 2020), 2018.

560   Straka, J. M.: Cloud and precipitation microphysics, Cambridge University Press, Cambridge., 2009.

Thompson, G., Field, P. R., Rasmussen, R. M. and Hall, W. D.: Explicit forecasts of winter precipitation using an improved
bulk microphysics scheme. Part II: Implementation of a new snow parameterization, Mon. Weather Rev., 136(12), 5095–5115,
doi:10.1175/2008MWR2387.1, 2008.

Thompson, P. D.: A transformation of stochastic equation for droplet coalescence, Bull. Am. Meteorol. Soc., 49(5), 596, 1968.

565   Ulbrich, C. W.: Natural variations in the analytical form of the raindrop size distribution, J. Clim. Appl. Meteorol., 22(10),
1764–1775, doi:10.1175/1520-0450(1983)022<1764:NVITAF>2.0.CO;2, 1983.