Stochastic modeling of particle structures in spray fluidized bed agglomeration using methods from machine learning

Lukas Fuchs*[1], Sabrina Weber*[1], Jialin Men[2], Niklas Eiermann[1], Orkun Furat[1], Andreas Bück[2], Volker Schmidt[1]

[1]Institute of Stochastics, Ulm University, Ulm, Germany

[2]Institute of Particle Technology, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany

* These authors contributed equally

E-mail adresse: lukas.fuchs@uni-ulm.de

Keywords: *Agglomeration, Copula, Random Forest, Spray Fluidized Bed*

Agglomeration is an industrially relevant process for the production of bulk materials in which the product properties depend on the morphology of the agglomerates, e.g., on the distribution of size and shape descriptors. Thus, accurate characterization and control of agglomerate morphologies is essential to ensure high and consistent product quality. This paper presents a pipeline for image-based inline agglomerate characterization and prediction of their time-dependent multivariate morphology distributions within a spray fluidized bed process with transparent glass beads as primary particles. The framework classifies observed objects in image data into three distinct morphological classes–primary particles, chain-like agglomerates and raspberry-like agglomerates–using various size and shape descriptors. To this end, a fast and robust random forest classifier is trained. Additionally, the fraction of primary particles belonging to each of these classes, either as individual primary particles or as part of a larger structure in the form of chain-like or raspberry-like agglomerates, is described using parametric regression functions. Finally, the temporal evolution of bivariate size and shape descriptor distributions of these classes is modeled using low-parametric regression functions and Archimedean copulas. This approach improves the understanding of agglomerate formation and allows the prediction of process kinetics, facilitating precise control over class fractions and morphology distributions.

1 Introduction

Agglomeration is a widely used particle formulation process that improves the handling of intermediate bulk solids, such as powder, pellets and granules that require further processing before their final application, and generates solid materials with desired end-user product features. In agglomeration processes, primary particles are combined into larger clusters (agglomerates) by establishing bonds between individual primary particles. These bonds can be established due to interaction forces (e.g., electrostatic or van der Waals forces), by chemical bonds [1, 2] due to surface reactions of the contacting particles, or by capillary forces due to liquid bridges or by solid bridges. In the latter two examples, a liquid or a solid-containing liquid is required to provide the material that generates the bridges between the primary particles. Agglomeration technologies that utilize these principles are, e.g., pressure agglomeration, binder agglomeration, spray agglomeration and thermal agglomeration [3]. Agglomerates can have superior properties compared to powders of primary particles, e.g., better flowability, higher bulk density and improved mechanical properties, etc., resulting in less dust formation and better strength and durability as well as a high surface-to-volume ratio [4, 5, 6]. The morphology, i.e., the structure of the agglomerates, determines the performance of the bulk material [7].

Spray fluidized bed (SFB) agglomeration is a common agglomeration technique that combines agglomeration and drying in a single vessel by atomizing a binding agent onto a bed of solid particles fluidized by hot gas. SFB agglomeration finds numerous applications in the chemical, food and pharmaceutical industry as it allows mixing, structure formation and drying of agglomerates in a single apparatus. This enables, e.g., the containment of dust while ensuring cost effectiveness and high efficiency [8, 9, 10, 11, 12]. These advantages are further enhanced if the spray agglomeration is operated in continuous mode, as this enables uniform operation with constant production rates and uniform agglomerate properties.

In situ process information on the agglomerate structure is required to control the agglomerate formation process with respect to the kinetics and towards autonomous process control. In situ information allows for inverse design of agglomerate structures, i.e., based on product requirements process conditions are selected such that the desired structure is achieved. Inverse design can be used to implement model-based feedback control that autonomously drives agglomerate formation processes to produce desirable agglomerate structures [13, 14, 15, 16, 17].

However, agglomeration typically occurs within a timescale of minutes. To enable process control, information on agglomerate formation must be obtained on smaller time scales, e.g., within several seconds. An option to fulfill this time constraint are sequences of in situ high-speed images of individual agglomerates, from which morphological descriptors can be extracted to

quantify the current state of the agglomeration process. This is routinely done with respect to agglomerate size and some measures of sphericity, more detailed descriptors have not been considered yet. Implementation of the analysis in a recursive fashion, i.e., updating available information by new measurement information, could enhance the real-time capability of structure assessment.

This paper presents a computationally efficient pipeline for image-based particle analysis designed for inline agglomeration characterization for perspective use in autonomous process control, integrating image segmentation, object (particle or agglomerate) classification and parametric modeling of object morphologies. Specifically, the pipeline employs fast, convolution-based denoising [18] combined with Otsu thresholding [19] to effectively extract individual objects from image data. A comprehensive set of morphological descriptors is then computed from these segmented objects in order to characterize the agglomeration status and to classify observed objects in image data into three distinct classes: (i) primary particles, which have not yet agglomerated or have broken from agglomerates again; (ii) chain-like agglomerates, consisting of a few primary particles aligned in a nearly linear configuration; and (iii) raspberry-like agglomerates, which denotes large clusters of agglomerated particles with multiple contact points between them. For fast classification, methods from artificial intelligence [20] are utilized to achieve high computational efficiency, suitable for inline classification. The classification enables the subsequent modeling of descriptor distributions for each individual particle/agglomerate classes, using parametric families of probability distributions. To gain an even deeper understanding of the state of agglomeration within the process, it is important to consider descriptors that capture both the size and shape of objects within each class. To account for the dependency between these descriptors, we use so-called Archimedean copulas to determine joint distributions of the considered descriptors [21, 22, 23]. Utilizing regression techniques on the parameters of these copula-based models enables time-dependent modeling and prediction of distributions of size and shape descriptor (vectors), providing insight into structural evolution of agglomerates over time [24]—an essential step towards model-based control, especially in the context of model-predictive control.

This paper is structured as follows. Section 2.1 describes the experimental study of agglomeration in the SFB using glass beads. Then, the imaging procedure and the subsequent image segmentation is explained in Section 2.2. Section 2.3 presents various geometrical descriptors used for classifying particles and agglomerates, with a random forest classifier explained in more detail in Section 2.4. Moreover, a parametric modeling approach for bivariate distributions of size and shape of particles/agglomerates as well as the temporal evolution of these distributions is explained in Section 2.5.2. This is followed by a sensitivity analysis that investigates the quantity of model quality for different amounts of data in Section 2.6. Section 3 provides the results of classification, time-dependent modeling and the sensitivity analysis. Section 4 concludes this contribution.

2 Materials and methods

2.1 Experimental setup

This work is based on the experimental work on SFB agglomeration of [25]. We briefly summarize the experimental setup and conditions that generate the agglomerates used in this study.

A pilot scale cylindrical fluidized bed with an inner diameter of 300 mm was used (general setup depicted in Figure 1). A two-fluid spray nozzle (Düsen-Schlick GmbH, model 940/6 with a hemispheric cap, liquid orifice diameter: 0.8 mm) was used, positioned 420 mm above the air distributor plate. The heated fluidization air enters through the air distributor plate. Primary particles are suspended by the fluidization gas. An aqueous binder solution is sprayed onto the fluidized particles, so that the surface of the particles is wetted and liquid bridges are formed after inter-particle collisions. Hot air dries and transforms the liquid bridges into solid bridges, thereby forming agglomerates. A filter is used to remove dust from the exhaust gas before it exits the equipment.

In the experiments [25], transparent glass beads are used as primary particles ($\rho_\mathrm{p} = 2500\,\mathrm{kg/m^3}$, average volume-weighted diameter $d_{1,3} = 0.24\,\mathrm{mm}$). Hydroxy-propyl-methyl-cellulose (Pharmacoat 606 from Shin-Etsu, Japan) is taken as binder solution (binder content reported in Table 1). The experiments were performed in continuous mode with fixed atomization flow rate ($0.08\,\mathrm{m^3/min}$) and average binder spray rate ($32\,\mathrm{g/min}$). Conditions for all experiments denoted by Experiments A - E [25] are summarized in Table 1. Experiments A, B and C were performed with the same binder content (4 wt-%) but different gas inlet temperatures which were set to 80 °C, 90 °C, 100 °C, respectively. Experiments D, B and E were conducted with same gas inlet temperature of 90 °C and the binder content in the spray was set to 2 wt-%, 4wt-% and 6 wt-%, respectively. All experiments ran for 120 minutes. Samples were taken from the bed and the product outlet at time steps $t \in T = \{10, \ldots, 120\}$. The present work will be based on data collected from their experiments A and E, having the most different experimental conditions (extreme cases).

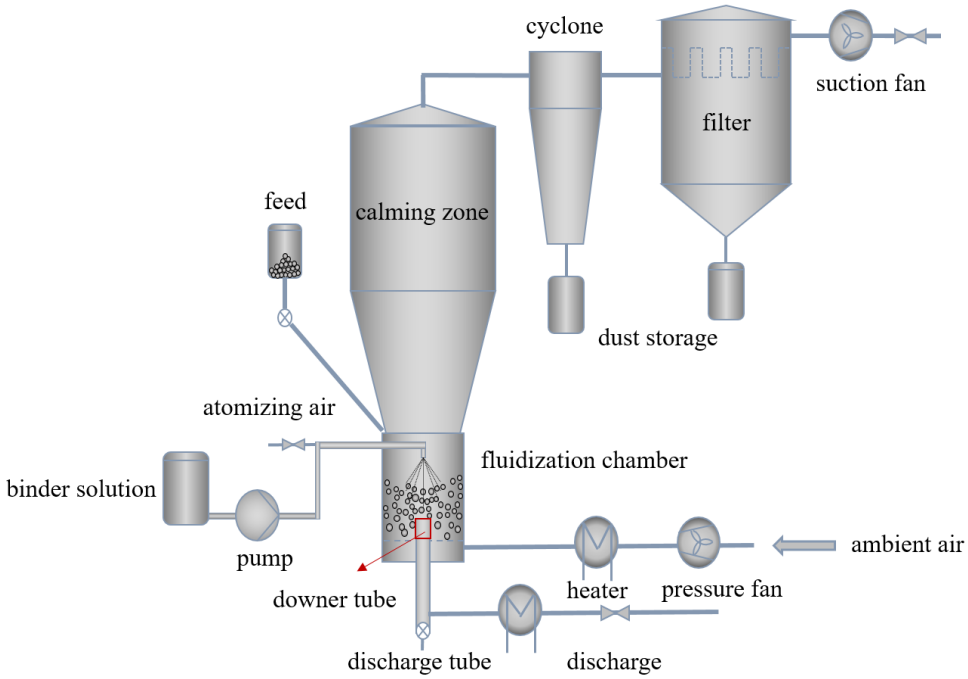

Figure 1: **Process setup.** Scheme of the pilot-scale spray fluidized bed for continuous spray agglomeration [25].

Table 1: **Process parameters.** Experimental conditions for spray fluidized bed agglomeration

	A	B	C	D	E
Inlet gas temperature [°C]	80	90	100	90	90
Binder content in wt%	4	4	4	2	6
Fluidization air mass flow rate [kg/hr]	284	286	280	282	280
Particle feed rate [g/min]	145.5	158.5	181.9	165.9	154.6

2.2 Imaging and image processing

In-situ image sequences of particles/agglomerates were acquired using commercial equipment (Camsizer, MicroTracRetsch). There, sample material of the experiment is putted into a dosage hopper. A vibrating chute then guides the particles so that they fall freely in front of an illuminated plane. During this free fall high-speed images are taken at 60 fps with an image size of 1012 pixels by 742 pixels at a spatial resolution of 15 µm/pixel. Per experiment and sampling time point, at least 20 images were acquired, with 21.6 objects being observed per image on average .

In order to characterize individual objects within image data and their descriptor distributions, the objects have to be extracted first from the images.

An image $I\colon \mathcal{X} \to \{0,\dots,255\}$ is considered as a mapping from the set of all pixel positions $\mathcal{X} = \{1,\dots,1012\} \times \{1,\dots,742\}$ to the set of 8-bit pixel values. To extract individual particles/agglomerates, first each pixel is classified as either background or foreground, i.e., a function $\bar{I}\colon \mathcal{X} \to \{0,1\}$ is computed to classify pixels. More precisely, a pixel $(i,j) \in \mathcal{X}$ with $\bar{I}(i,j) = 1$ is considered to be a foreground pixel, whereas $\bar{I}(i,j) = 0$ indicates a background pixel. The function $\bar{I}$ is computed by utilizing a non-local means filter [18] followed by an Otsu thresholding [19]. The former decreases noise present in an image I, whereas the latter is a histogram based heuristic for threshold computation. More precisely, by means of non-local denoising a smoothed version $I'\colon \mathcal{X} \to [0,255]$ of $I\colon \mathcal{X} \to \{0,255\}$ is computed which is given by

$$I'(x) = \frac{\sum\limits_{z \in Z_x^{21}} w(x, x+z) \cdot I(x+z)}{\sum\limits_{z \in Z_x^{21}} w(x, x+z)}, \qquad \text{with} \qquad Z_x^{21} = \{z \in \mathbb{Z}^2 \colon |z|_\infty \leq 21, x+z \in \mathcal{X}\}, \tag{1}$$

where $w(x,y)$ is a distance-based weight, given by

$$w(x,y) = \exp\left(-\frac{1}{100} \sum_{z \in Z_{x,y}^5} \left(I(x+z) - I(y+z)\right)^2\right), \qquad \text{with} \qquad Z_{x,y}^5 = \{z \in \mathbb{Z}^2 \colon |z|_\infty \leq 5, x+z, y+z \in \mathcal{X}\}. \tag{2}$$

Here, $|\cdot|_\infty$ denotes the maximum norm and $\mathbb{Z} = \{\ldots, -1, 0, 1, \ldots\}$ the set of all integers. Intuitively, the weight $w(x, y)$ measures the similarity between the intensity patterns (so-called patches) surrounding the pixels x and y. Pixels with more similar local patterns (in terms of the squared sum of their difference) contribute more significantly to the value in the denoised image.

After computing the smoothed image I', Otsu's method is applied to determine a global intensity threshold $\eta \in [0, 255]$, see [19] for detail. By applying the global threshold η to I', this yields the phasewise segmentation $\bar{I}\colon \mathcal{X} \to \{0, 1\}$, given by

$$\bar{I}(x) = \begin{cases} 0, & \text{if } I'(x) < \eta, \\ 1, & \text{else.} \end{cases} \tag{3}$$

An object-wise segmentation $p\colon \mathcal{X} \to \mathbb{N} \cup \{0\}$ is subsequently achieved by setting $p(x) = 0$ if $\bar{I}(x) = 0$ and assigning each connected components of pixel positions $x \in \mathcal{X}$ with $\bar{I}(x) = 1$ a unique positive integer. This value is often referred to as the index (or label) of the connected component. Thereby, two pixel positions $x, y \in \mathcal{X}$ are considered to belong to the same connected component if and only if there exist a sequence $z_1, \ldots, z_k \in \{x \in \mathcal{X}\colon \bar{I}(x) = 1\}$ such that $x = z_1, z_k = y$ and $|z_i - z_{i+1}|_2 = 1$ for all $i \in \{1, \ldots, k-1\}$, where $|\cdot|_2$ is the Euclidean norm. In the following, pixel positions belonging to the object with index i are denoted by $p_i = \{x \in \mathcal{X}\colon p(x) = i\}$.

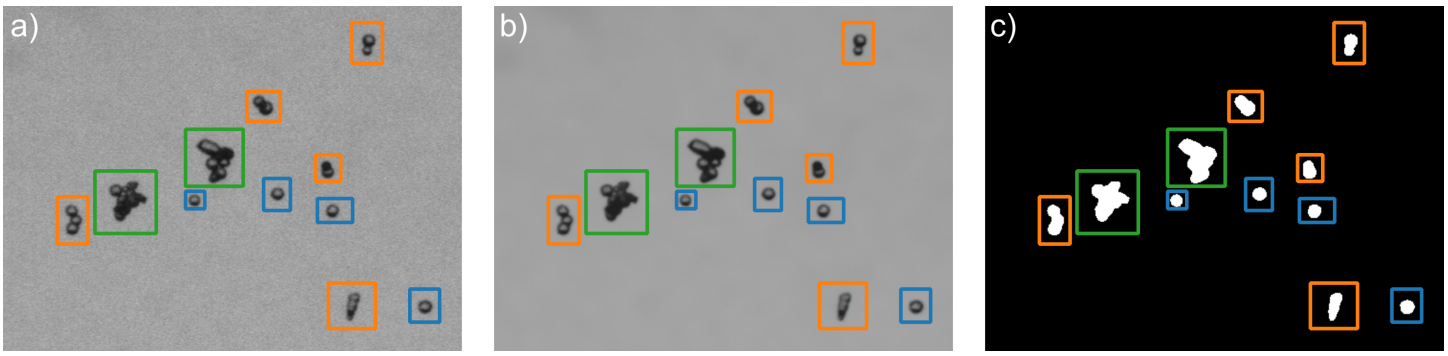

Figure 2: **Preprocessing of images.** a) Exemplary cutout of an image depicting particles/agglomerates of Experiment A at time step 50 min. b) Denoised (non-local means) of a). c) Corresponding phasewise segmentation of b). Exemplary, primary particles are highlighted in blue, chain-like agglomerates in orange and raspberry-like agglomerates in green.

2.3 Structure analysis and particle classes

During the SFB agglomeration process, the primary particles agglomerate. Our aim is to characterize the size and shape of both the agglomerates and the primary particles. To quantify the state of agglomeration, we divide the objects observed in image data into three different classes: primary particles, chain-like agglomerates and raspberry-like agglomerates.

To decide the class membership of objects observed in image data, we use various geometrical descriptors such as the area-equivalent diameter, which is given by $d(p_i) = 2\sqrt{a(p_i)/\pi}$ for particle/agglomerate with index i, where $a(p_i)$ is the area of p_i. Note that we compute the area of the i-th particle/agglomerate p_i by deploying the point-count method [26]. A further geometrical descriptor considered is the area $a_{\text{convex}}(p_i)$ of the convex hull of p_i, where again the point-count method is deployed for the computation of the area. Figure 3a) visualizes the difference between the area of p_i and its convex hull in red.

Then, the solidity of p_i is given by $s(p_i) = a(p_i)/a_{\text{convex}}(p_i)$. The solidity is a geometrical descriptor that quantifies how much the shape of p_i deviates from being perfectly convex. A solidity value of 1 indicates a fully convex object. Additionally, we compute the lengths of the major and minor axes of the ellipse that has the same normalized second central moments as p_i, see Figure 3b). Further details on the computation of such a "moment-equivalent" ellipse can be found in [27]. The lengths of the major and minor axes are denoted by $v_1(p_i), v_2(p_i) \in [0, \infty)$, respectively. Based on these lengths, the eccentricity $e(p_i)$ is computed by

$$e(p_i) = \sqrt{1 - \left(\frac{v_2(p_i)}{v_1(p_i)}\right)^2}, \tag{4}$$

see [28]. The eccentricity can be interpreted as a measure that quantifies how much a particle/agglomerate deviates from the circular shape, where a circle has an eccentricity of 0. The orientation $o(p_i) \in [-\pi/2, \pi/2)$ is computed as the angle between the major axis and the y-axis of the coordinate system, see Figure 3c). Figure 3d), shows the minimal radius $r_1(p_i)$ of the sphere that encloses the object and the maximum radius $r_2(p_i)$ of the sphere that inscribes the object. The relationship between these two descriptors, i.e., $r(p_i) = r_1(p_i)/r_2(p_i)$, is a further descriptor that is considered for classifying objects.

Additionally, we consider Feret diameters to construct further geometrical descriptors. The Feret diameter into a direction is defined as the distance between two parallel planes that are normal with respect to the chosen direction and enclose the convex hull of an object [29, 30]. A visualization of the Feret diameter of an object in an exemplarily chosen direction is given in Figure 3e). For classification purposes, we use the largest Feret diameter of p_i as an additional geometrical descriptor, which we denote by $h(p_i)$.

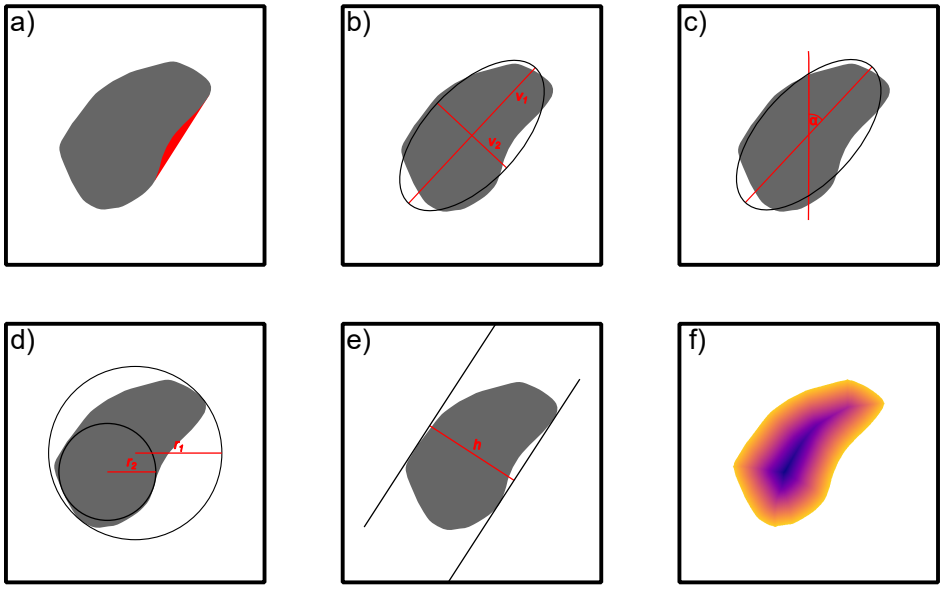

Figure 3: **Illustration of geometrical descriptors.** (a): The convexity is computed by dividing the area of the object by the area of the convex hull (red). (b): An ellipsoid is fitted to the object and the major and minor axis are determined. (c): The orientation α is given by the angle between the major axis of a fitted ellipsoid and a vertical line. (d): maximum radius of sphere that inscribe the object and minimum radius of sphere that enclose the object. (e): distance of each pixel to the object border. (f): One exemplary Feret diameter h.

Further geometrical descriptors of p_i are computed from the Euclidean distance transform of p_i which assigns each pixel of p_i with the pixel's Euclidean distance to the boundary of p_i, see Figure 3f). Then, we compute the mean and standard deviations of the Euclidean distances of all pixels associated with p_i, denoted by $\lambda_{\text{mean}}(p_i)$ and $\lambda_{\text{std}}(p_i)$, respectively. In addition, we compute the maximum Euclidean distance of pixels to the particle border $\lambda_{\max}(p_i)$ as a geometrical descriptor for classification purposes. Moreover, for each p_i, we compute the perimeter $\delta(p_i)$ by identifying boundary pixels using four-neighborhood connectivity. The length of each boundary pixel is then determined based on its local neighborhood configuration: a pixel with two vertically or horizontally adjacent neighbors is assigned a length of 1, a pixel with two diagonally adjacent neighbors is assigned a length of $\sqrt{2}$, and a pixel with one vertically or horizontally adjacent neighbor and one diagonally adjacent neighbor is assigned a length of $\frac{1+\sqrt{2}}{2}$. The total perimeter $\delta(p_i)$ is obtained by summing these lengths and multiplying by the pixel length [31] is used. With the perimeter and the area of object i, we can compute the roundness $\psi(p_i)$ in the following way, see [30]:

$$\psi(p_i) = 4\pi \frac{a(p_i)}{\delta(p_i)^2}. \tag{5}$$

For classification purposes, we also consider the centroid and the gray value-weighted centroid of p_i, denoted as z_1 and z_2, respectively. The latter is calculated by weighting each pixel's contribution according to its gray value. Then, the Euclidean distance between the centroid and the weighted centroid, i.e., $z(p_i) = |z_1(p_i) - z_2(p_i)|$, is considered as additional geometrical descriptor.

Another descriptor to characterize object i is the fractal dimension $\kappa(p_i)$, which is computed using the tiled box counting method, see [32, 33]. Last but not least, the texture of an object is characterized by computing descriptors that quantify the gray values of pixels associated with objects.

Specifically, we compute the following textural descriptors: the mean gray value $g_{\text{mean}}(p_i)$ of pixels associated with p_i, the standard deviation $g_{\text{std}}(p_i)$ of gray values of pixels associated with p_i, as well as the minimum $g_{\min}(p_i)$ and maximum $g_{\max}(p_i)$ gray value of pixels associated with p_i. In total, this work utilizes $\nu = 22$ different descriptors for classification purposes. For a better overview, the descriptors are summarized in Table 2.

Table 2: **Overview of geometrical and textural descriptors for particles.**

Geometrical/textural descriptor	Symbol	Range	Unit
area-equivalent diameter	d	$(0, \infty)$	µm
area	a	$(0, \infty)$	µm^2
area of convex hull	a_{convex}	$(0, \infty)$	µm^2
solidity	s	$(0, 1]$	-
length of major axis	v_1	$(0, \infty)$	µm
length of minor axis	v_2	$[0, \infty)$	µm
eccentricity	e	$[0, 1]$	-
orientation	o	$[-\pi/2, \pi/2)$	-
minimal radius of a sphere enclosing the object	r_1	$(0, \infty)$	µm
maximum radius of a sphere inscribing the object	r_2	$(0, \infty)$	µm
ratio of r_1 and r_2	r	$(0, \infty)$	-
largest Feret diameter	h	$(0, \infty]$	µm
mean border distance	λ_{mean}	$(0, \infty)$	µm
standard deviation border distance	λ_{std}	$[0, \infty)$	µm
maximal border distance	λ_{max}	$(0, \infty)$	µm
perimeter	δ	$[0, \infty)$	µm
roundness	ψ	$[0, \infty)$	-
centroid	z_1	$[0, 15.3] \times [0, 11.13]$	mm
gray-value weighted centroid	z_2	$[0, 15.3] \times [0, 11.13]$	mm
fractal dimension	κ	$[0, 2]$	-
mean gray value	g_{mean}	$[0, 255]$	-
standard deviation gray values	g_{std}	$[0, 255]$	-
minimum gray value	g_{min}	$[0, 255]$	-
maximum gray value	g_{max}	$[0, 255]$	-

2.4 Particle classification

In this section we describe the procedure to assign each object in the segmented image data to one of the following classes: primary particles, chain-like agglomerates and raspberry-like agglomerates. Since hand-labeling is impractical for large datasets, especially in online image processing, and no direct functional relationship is known between geometrical/textural descriptors (as introduced in Section 2.3) and object classes, a classification tool from artificial intelligence is employed. Specifically, a commonly used random forest classification framework [20] is trained to learn a mapping $m\colon \mathbb{R}^\nu \to \{0, 1, 2\}$ from the set of interpretable descriptor vectors to a class label, where the label 0 corresponds to primary particles, 1 to chain-like agglomerates, and 2 to raspberry-like agglomerates.

Roughly speaking, a random forest consists of several decision trees that will be used to generate a "vote" on the class membership of a particle/agglomerate with descriptor vector $P \in \mathbb{R}^\nu$. The random forest's classification is given by the majority of the votes of the individual trees.

Thus, in order to introduce the notion of a random forest, we first introduce binary decision trees. We will denote a binary decision tree $T = (V, E, \mathcal{F}, w)$ as a quadruple of a set of vertices V, a set of edges $E \subset V \times V$, a set of decision functions $\mathcal{F}$ and a function of w that implicitly assigns a class label to P. The graph (V, E) forms a perfect binary tree, meaning that each node $v \in V$ has either zero or exactly two children. The set of children of v is denoted by $N(v) = \{v' \in V \mid (v, v') \in E\}$. The tree has a unique root node $v_\text{r} \in V$, and all leaf nodes $V_\text{l} = \{v \in V\colon N(v) = \emptyset\}$ are at the same depth. The depth of a leaf node $v_\text{l} \in V_\text{l}$ is given by the length k of the shortest sequence $e_1, \ldots, e_k \subset E$ in which consecutive edges share a node, and v_r, v_l are contained in e_1, e_k, respectively. The set of decision functions $\mathcal{F} = \{f_v\colon \mathbb{R}^\nu \to N(v)\colon v \in V, N(v) \neq \emptyset\}$ contains for each non-leaf node a function that maps each descriptor vector $P \in \mathbb{R}^\nu$ to a unique child note $u \in N(v)$. Furthermore, the function $w\colon V_\text{l} \to \{0, 1, 2\}$ maps each leaf node to a particle class in $\{0, 1, 2\}$. The vote of the decision tree T for a descriptor vector P is determined as follows: By computing the unique path $(v_\text{r}, f_v(P), f_{f_v(P)}(P), \ldots, v_\text{l}) \subset V$ from the root node v_r of the tree T to a leaf node $v_\text{l} \in V_\text{l}$ is identified. Then, the tree T assigns the object with the descriptor vector P to the class with label $w(v_\text{l})$.

In this work, the decision functions $f_v \in \mathcal{F}$ are threshold-based decisions applied to individual descriptors. Specifically, the decision functions take the form:

$$f_v(P) = \begin{cases} v_1, & \text{if } P_j > P^*, \\ v_2, & \text{else,} \end{cases} \tag{6}$$

where $\{v_1, v_2\} = N(v)$ are the child nodes of v, $P^* \in \mathbb{R}$ is a threshold and P_j is the j-th descriptor within the descriptor vector $P = (P_1, \dots, P_\nu)$. Thus, the decision function f_v is uniquely determined by the tuple (j, P^*) of the considered descriptor given by $j \in \{1, \dots, \nu\}$ of the descriptor vector $P \in \mathbb{R}^\nu$ and the corresponding threshold P^*.

For training data consisting of a sequence $\mathcal{P} \subset \mathbb{R}^\nu$ of descriptor vectors and corresponding class labels in $\{0, 1, 2\}$ for each descriptor vector, a decision tree T can be efficiently computed by the so-called CART algorithm [20]. Specifically, for a perfect binary tree (V, E) this algorithm, iteratively, determines the functions $f_v \colon \mathbb{R}^\nu \to N(v)$ for all $v \in V$ with $N(v) \neq \emptyset$ by greedily maximizing some quality measure $Q(f_v, \mathcal{P})$ for some set of descriptor vectors $\mathcal{P}$.

We consider a quality measure that is based on the so-called Gini coefficient G. The Gini coefficient gives a measure for the impurity of classes of some sequence $\mathcal{P}$ of descriptor vectors. More precisely, this coefficient is given by

$$G(\mathcal{P}) = \sum_{k=0}^{2} (1 - q_k) q_k, \tag{7}$$

where $q_k \in [0, 1]$ is the fraction of descriptor vectors of $\mathcal{P}$ that were assigned to the class k. The quality $Q(f_v, \mathcal{P})$ of a function f_v for a sequence $\mathcal{P} \subset \mathbb{R}^\nu$ of descriptor vectors is then given by

$$Q(f_v, \mathcal{P}) = \sum_{u \in N(v)} G(\mathcal{P}_u) |\mathcal{P}_u|, \tag{8}$$

where $\mathcal{P}_u$ is the maximum subsequence of P for which it holds that $f_v(P) = u$ for all $v \in \mathcal{P}_u$, and $|\cdot|$ denotes the length of the sequence under consideration.

The CART-algorithm starts with the root $v_\mathrm{r} \in V$ of a tree (V, E) and the sequence $\mathcal{P}$ of all measured particle descriptors, and computes a function f_{v_r} that minimizes $Q(f_{v_\mathrm{r}}, \mathcal{P})$. Afterward, the same procedures are repeated for child nodes $u \in N(v_\mathrm{r})$ with the corresponding sequences $\mathcal{P}_u$ of descriptor vectors until, for all non-leaf nodes $v \in V$, the function f_v is determined. At a last step, the value of the function $w \colon \{v \in V \colon N(v) = \emptyset\} \to \{0, 1, 2\}$ is set to the class in $\{0, 1, 2\}$ that is most frequently among the classes of the particle descriptor vectors of $\mathcal{P}_v$ in the training data. See [20] for more details.

This procedure results in an optimal classification of the training data [20]; however, considering all components $P_j,\ j \in \{1, \dots, \nu\}$ of the descriptor vectors $P \in \mathcal{P}$ in the computation of $f_v,\ v \in V,\ N(v) = \emptyset$ often leads to overfitting, and thus poor performance on unseen data. To mitigate this, a subset $J \subset \{1, \dots, \nu\}$ of $|J| < \nu$ descriptor component indices is chosen before training a decision tree T. Then, when computing, the decision functions $\mathcal{F}$ are restricted to descriptors with an index in J.

A random forest is a collection of $B \in \mathbb{N}$ binary decision trees $T_1, \dots, T_B$. These are constructed by independently calibrating $B > 1$ binary decision trees $T_1, \dots, T_B$, each of which with an independently chosen random subset $J_1, \dots, J_B \subset \{1, \dots, \nu\}$ of descriptor indices. Then, the random forest can be deployed for classifying an object with descriptor vector $P \in \mathbb{R}^\nu$, by (i) determining the B class assignments of P according to decision trees $T_1, \dots, T_B$ followed by (ii) majority voting along these class assignments. This approach effectively addresses the overfitting problem. The choice of a random tree's hyperparameters—the number B of considered decision trees as well as the number $|J|$ of considered particle descriptors, and the depth of the binary trees $(V_1, E_1), \dots, (V_B, E_B)$—is subject to the training procedure. For more details and heuristics for choosing these hyperparameters, see [34]. We optimize these hyperparameters, by means of a grid-search. The grid-search is performed by training and evaluating random forests with 50, 100 and 200 decision trees, considering maximum tree depths of 3, 5, 7 and ν, and sizes of J equal to $5 \approx \sqrt{\nu}$ and ν. During the gird-search, all possible combinations of hyperparameters are used for training. The best hyperparameters are those that lead to the random forest with the highest percent of correctly classified particles/agglomerates on the training data. The prediction of a class label using a random forest, is very fast, thus combined with the fast computation of the image segmentation and particle descriptors, an online classification of imaged particles/agglomerates is enabled. Furthermore, the decisions f_v of the individual trees can be interpreted, giving insight into, the influence of interpretable particle descriptors on the classification, see Section 3.2.

2.5 Particle class and descriptor modeling

The previously introduced automatic classification method enables the efficient classification of a large particle/agglomerate database. In this section, the focus is on modeling the evolution of primary particles, chain-like and raspberry-like agglomerates over time. This is achieved by analyzing two aspects: first, the temporal evolution of the class size fractions, and second, the evolution of bivariate distributions of descriptors within individual classes over time. In the following, methods for modeling these evolutions are presented generally. Later, these methods will be deployed to data sets acquired from the Experiments A and E, see Section 3 for further details.

2.5.1 Temporal evolution of class sizes

First, we analyze the size of an object class over time by means of the area-weighted fraction of objects belonging to this class compared to all observations. For this, let M be a set of pairs (t, y) of time steps $t \in T$ and corresponding size fractions

$y \in [0,1]$ of the considered class. We will model the value of y with respect to t by means of a parametric regression function $\zeta\colon [0,\infty) \to \mathbb{R}$. The specific form of the regression functions is given by

$$\zeta(t) = c_1 - c_2 \exp(-c_3\, t) \tag{9}$$

for any $t \in [0,120]$, where $c_1, c_2, c_3 \in \mathbb{R}$ are the parameters of the regression function and $\exp\colon \mathbb{R} \to [0,\infty)$ is the Euler's function. Thereby, the parameter c_1 describes the asymptotic value of g for $t \to \infty$; c_2 determines the value of g for small values of t, and c_3 determines how fast g reaches its asymptotic value c_1.

For a set M of pairs (t, y) the parameters of the regression function ζ can be calibrated by minimizing the mean squared error (MSE), i.e., the values of $(c_1, c_2, c_3) \in \mathbb{R}^3$ are given by

$$(c_1, c_2, c_3) = \underset{(c_1,c_2,c_3)\in\mathbb{R}^3}{\operatorname{argmin}} \frac{1}{|M|} \sum_{(t,y)\in M} (\zeta(t) - y)^2. \tag{10}$$

For both experiments, Experiment A and Experiment E, we use the regression function, given by Equation (9), to model the size fractions of the primary particles, and to model the size fractions of the raspberry-like agglomerates over time. Thus in these four cases, the set M consist of pairs (t, y) of a time steps $t \in T$ and the respective area-weighted size fractions of primary particles (or raspberry-like agglomerates) at time t at all respective experiments. In order to ensure that the size fractions, modeled by means of the resulting regression functions, are fractions, the evolution of the size fraction of the chain-like agglomerate class is modeled by $1 - g_1 - g_2$, where g_1, g_2 are the respective regression functions of the primary particles and raspberry-like agglomerates. Although there exist values of $(c_1, c_2, c_3) \in \mathbb{R}^3$ for which $g_1(t), g_2(t) \notin [0,1]$ for $t \in [10,120]$, this does not occur in the application, see Section 3.3.1. The choice of modeling the size fraction of the chain-like agglomerates as the complement of the other two classes is intuitive, since chain-like agglomerates appear as an intermediate class between non-agglomerated primary particles and the desired final product, the raspberry-like agglomerates. In Section 3 the results of the regressions for both considered experiments and all classes of objects are shown.

2.5.2 Parametric copula-based modeling

To gain a more comprehensive understanding of the agglomeration process, the bivariate probability distributions of area-equivalent diameter d and solidity s within the three classes are modeled parametrically for each time step. This parametric modeling facilitates the time-dependent regression of probability densities by performing the regression in a lower dimensional space instead, namely, on the set of model parameters. Consequently, it enables the regression of particle descriptor distributions and thus allows for predicting these distributions for time steps which were not directly measured. In this manner a temporal model for the bivariate probability distribution of descriptor vectors can be obtained. This modeling approach is deployed for each class and observed in Experiments A and E individually, see Section 3. However, to facilitate the notation the methodology is introduced in general for one and two-dimensional vector data in this section. Therefore, unless stated otherwise, we consider a single class of objects observed in a single experiment in the remainder of Section 2.5.2.

Univariate densities. First, we want to determine the univariate densities $f_d, f_s : \mathbb{R} \to [0,\infty)$ and the corresponding cumulative distribution functions $F_d, F_s : \mathbb{R} \to [0,1]$ of d and s for an individual time step. Therefore, we use the parametric families $\mathcal{G} = \{\text{normal}, \text{log-normal}, \text{gamma}\}$, which are further specified in Table 3 and in [35]. However, note that the solidity takes values in $[0,1]$ and the area-equivalent diameter takes values in $(0,\infty)$. Thus, for parametric families with different support we consider truncated versions to ensure a correct support. To determine suitable parameter values and the best family, maximum likelihood estimation is used [36]. In this manner, we obtain for each family $G_d, G_s \in \mathcal{G}$ parametric probability densities $f_d^{G_d,\omega_{t,d}}, f_s^{G_s,\omega_{t,s}} : \mathbb{R} \to [0,\infty)$ of d and s at the considered time step t as well as their corresponding distribution functions $F_d^{G_d,\omega_{t,d}}, F_s^{G_s,\omega_{t,s}} : \mathbb{R} \to [0,1]$. Note that $\Omega_{G_d}, \Omega_{G_s}$ denotes the parameter space of family G_d and G_s. The used parameter vector for d and s are denoted by $\omega_{t,d} \in \Omega_{G_d}$ and $\omega_{t,s} \in \Omega_{G_s}$ denotes.

More details on fitting these parameters vector and determining the overall best family for all time steps is given in subsequent paragraphs.

Table 3: **Parametric univariate distributions.** Parametric families of univariate distributions with corresponding density, support and parameter space Ω, where Γ denotes the gamma function [35].

parametric family	probability density	support	$\omega \in \Omega$
normal	$\frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/(2\sigma^2)}$	$(-\infty,\infty)$	$(\mu,\sigma) \in \mathbb{R} \times (0,\infty)$
log-normal	$\frac{1}{x\sqrt{2\pi\sigma^2}} e^{-(\log^2(x)-\log^2(\mu))/(2\sigma^2)}$	$(0,\infty)$	$(\mu,\sigma) \in (0,\infty) \times (0,\infty)$
gamma	$\frac{1}{\beta^\alpha} x^{\alpha-1} e^{\frac{1}{\beta}x} \frac{1}{\Gamma(\alpha)}$	$[0,\infty)$	$(\alpha,\beta) \in (0,\infty) \times (0,\infty)$

Bivariate copula-based densities. To describe the structure of objects at an individual time step, the interdependence between the size and shape descriptors can be modeled. To do so, the bivariate distribution of solidity and area-equivalent diameter is modeled using copulas, which provide a flexible approach by decoupling the marginal distributions from their dependence structure. More precisely, a function $C : [0,1]^2 \to [0,1]$ is called a bivariate copula if C is the cumulative distribution function of a two-dimensional random vector with standard uniformly distributed marginals. Then, according to Sklar's representation formula (see [21]), for the joint cumulative distribution function $F_{d,s} : \mathbb{R}^2 \to [0,1]$ of d and s, there exists a copula C such that

$$F_{d,s}(x_d, x_s) = C(F_d(x_d), F_s(x_s)), \tag{11}$$

for $x_d \in [0,\infty)$ and $x_s \in [0,1]$ denoting the particle descriptor values.

Note that, assuming that $F_{\mathrm{s,d}}$ and C are differentiable, it follows from Equation (11) that the joint probability density $f_{d,s} : \mathbb{R}^2 \to [0,\infty)$ of d and s is given by

$$f_{d,s}(x_d, x_s) = c(F_d(x_d), F_s(x_s)) f_d(x_d)\, f_s(x_s), \tag{12}$$

where $c : [0,1]^2 \to [0,\infty)$ is the probability density of C. Thus, in this case, the differential version of Sklar's representation formula given in Equation (12) can be used to construct bivariate densities by fitting the univariate margins (see the preceding paragraph) and then a bivariate copula.

For modeling bivariate distributions, so-called Archimedean copulas are used in the present paper. The definition of these copulas is based on Archimedean generators $\varphi : [0,1] \to [0,\infty)$, which are continuous, strictly decreasing functions such that $\varphi(1) = 0$. Moreover, let $\varphi^{[-1]} : [0,\infty) \to [0,1]$ be the pseudo inverse of φ, i.e., $\varphi^{[-1]}(x) = \varphi^{-1}(x)$ if $0 \le x \le \varphi(0)$ and $\varphi(x) = 0$ if $x \ge \varphi(0)$, where φ^{-1} denotes the inverse of φ. An Archimedean copula is then given by

$$C(u_1, u_2) = \varphi^{[-1]}(\varphi(u_1) + \varphi(u_2)), \tag{13}$$

for any $u_1, u_2 \in [0,1]$, see e.g. [21]. To model bivariate densities, various parametric families $\{\varphi_\theta : \theta \in \Theta\}$ of Archimedean generators are considered, see Table 4. The space of admissible parameters is denoted by $\Theta \subset \mathbb{R}$. Each family of Archimedean generators leads to a parametric family of copula densities $\{c_\theta : \theta \in \Theta\}$ given by

$$c_\theta(u_1, u_2) = \frac{\partial^2}{\partial u_1 \partial u_2} \varphi_\theta^{[-1]}(\varphi(u_1) + \varphi(u_2)) \tag{14}$$

for any $u_1, u_2 \in [0,1]$ and $\theta \in \Theta$.

Even further families of copula densities can be constructed, by considering rotating copula densities within a given family by multiples of 90°. More precisely, c_θ can be rotated around the midpoint $(0.5, 0.5)$ by $90°, 180°$ or $270°$ to obtain copula families. To determine the optimal copula family and density parameter, maximum likelihood estimation is used as in the case of fitting the univariate distributions.

For the parametric case with Archimedean copulas we adapt Equation (12) in the following. Therefore, we consider a set $\mathcal{Z} = \{\text{Frank, Joe, Clayton, Gumbel, Ali-Mikhail-Haq}\}$ of copula types, each of which induces a parametric family of copula densities. The bivariate density with previously fitted marginal distributions and the copula density $c_{\mathrm{d,s}}^{Z,\theta_t} : [0,1]^2 \to \mathbb{R}$ of family Z with parameter $\theta_t \in \Theta_Z \subset \mathbb{R}$, where Θ_Z is the parameter space of the copula family Z defined in Table 4, is then given by

$$f_{d,s}^{Z,\theta_t}(x_d, x_s) = c_{\mathrm{d,s}}^{Z,\theta_t}(F_d^{G_d,\omega_{t,d}}(x_d), F_s^{G_s,\omega_{t,s}}(x_s)) f_d^{G_d,\omega_{t,d}}(x_d) f_s^{G_s,\omega_{t,s}}(x_s), \tag{15}$$

for any $x_d \in [0,\infty)$ and $x_s \in [0,1]$.

Table 4: **Archimedean generators.** Parametric families $\{\phi_\theta : \theta \in \Theta\}$ of Archimedean generators, together with their set of parameters $\Theta \subset \mathbb{R}$.

copula	Frank	Joe	Clayton	Gumbel	Ali-Mikhail-Haq
$\varphi_\theta(u)$	$-\ln \frac{\exp(-\theta u)-1}{\exp(-\theta)-1}$	$-\ln(1-(1-u)^\theta)$	$\frac{1}{\theta}(u^{-\theta}-1)$	$(-\ln u)^\theta$	$\ln \frac{1-\theta(1-u)}{u}$
Θ	$\mathbb{R} \setminus \{0\}$	$[1,\infty)$	$(0,\infty)$	$[1,\infty)$	$[-1,1)$

Time-dependent regression of distribution parameters. For our purpose of a time-dependent regression, it has to be ensured that the same parametric family of probability densities is deployed to model the distribution of a descriptor (vector) within a class for all time steps $T = \{10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120\}$ minutes of an experiment. Recall that, we

explain the procedure for a single class of objects observed in one single experiment to simplify the notation. For example, we can consider raspberry-like agglomerates in Experiment A. In all other cases the procedure is analog. We consider first the set $\mathcal{G} = \{\text{norm, lognorm, gamma}\}$ of distribution types, i.e., each type in $\mathcal{G}$ has an associated parametric family of univariate probability densities. In order to identify a parametric family of probability densities that can adequately model the the probability density of d for all time steps, we identify the optimal distribution type $\widehat{G}_d \in \mathcal{G}$ that maximizes the overall likelihood, i.e., the distribution type is given by

$$\widehat{G}_d = \arg\max_{G \in \mathcal{G}} \sum_{t \in T} \left(\max_{\omega_{t,d} \in \Omega_G} \sum_{x \in D_{t,d}} \log(f_d^{G,\omega_{t,d}}(x)), \right), \tag{16}$$

where $D_{t,d}$ denotes the set of descriptors d observed at time step t. The parametric family of univariate probability densities of the chosen distribution type $\widehat{G}_d$ is then fitted to the data $D_{t,d}$ for each time step by means of maximum likelihood estimation. Thus, we obtain a sequence of fitted parameters $\widehat{\omega}_{t,d} \in \Omega_{\widehat{G}_d}$ for each time step $t \in T$. In other words, for each $t \in T$ we obtain a sequence of univariate probability densities $f_d^{\widehat{G}_d,\widehat{\omega}_{t,d}}$ of d that all stem from the same parametric family. Analogously, univariate probability densities $f_s^{\widehat{G}_s,\widehat{\omega}_{t,s}}$ can be determined for modeling the distribution of the solidity for each time step $t \in T$ with the same parametric family.

After fitting the univariate distributions of both descriptors, we determine the best parametric family of copula densities by using Equation (15). More precisley, to chose the best-fitting copula family, the maximized likelihood values for the copula with previously fitted marginal distributions are summed up as above to identify the best-fitting family. More precisely, the best copula is provided by

$$\widehat{Z} = \arg\max_{Z \in \mathcal{Z}} \sum_{t \in T} \left(\max_{\theta_t \in \Theta_Z} \sum_{(x_d,x_s) \in D_{t,(d,s)}} \log(f_{d,s}^{Z,\theta_t}(x_d, x_s)) \right), \tag{17}$$

where $D_{t,(\mathrm{d,\ s})}$ is the set of two-dimensional descriptor vectors (pairs of area-equivalent diameter and the solidity) measured at time $t \in T$. Moreover, we obtain a sequence of fitted parameters $\widehat{\theta}_t \in \Theta_{\widehat{Z}}$ for each time step $t \in T$, i.e., we obtain a sequence of copulas $c_{\mathrm{d,s}}^{\widehat{Z},\widehat{\theta}_t}$, which together with the univariate distributions define the bivariate density $f_{\mathrm{d,s}}^{\widehat{Z},\widehat{\theta}_t}$, see Equation (15).

After this procedure we have identified a suitable low-parametric model for describing the temporally resolved bivariate distribution of area-equivalent diameters and solidity and the parameters, i.e., the vector $\tau_t = (\widehat{\omega}_{t,\mathrm{d}}, \widehat{\omega}_{t,\mathrm{s}}, \widehat{\theta}_t) \in \Omega_{\widehat{G}_\mathrm{d}} \times \Omega_{\widehat{G}_\mathrm{s}} \times \Theta_{\widehat{Z}}$, which is fitted to the set of two-dimensional descriptor vectors for each particle class, experiment and time step t.

At this point we are able to describe the bivariate distribution of area-equivalent diameter and the solidity for all particle classes and experiments at $t \in T = \{10, 20, \ldots, 120\}$. However, it is of interest to predict these distributions for all $t \in [10, 120]$. For this purpose, a regression of the parameters of the copula model is utilized. More precisely, the regression function of Equation (9) is utilized to predict the vector τ_t for all $t \in [10, 120]$. Note that, for each time step a different number of observed agglomerates is available. In order to always weight each available data point equally, the regression curve is fitted by minimizing the weighted MSE. Thus, Equation (10) is adjusted as follows

$$(c_1, c_2, c_3) = \underset{(c_1,c_2,c_3) \in \mathbb{R}^3}{\operatorname{argmin}} \sum_{t \in T} ((\tau_t^i - \ \zeta(t))|D_t|)^2, \tag{18}$$

where τ_t^i is the i-th entry of τ_t and $|D_t|$ is the amount of measured data at time step t . By utilizing a regression of the five parameter values over time by the described procedure, statements can be made about the distributions of unobserved points in time. Moreover, fitting functions to the parameter course over time allow to make predictive statements on the parametric distributions. Note that the described method is applied for Experiments A and E as well as for each particle class, i.e., primary particle, chain-like agglomerates and raspberry-like agglomerates. The results are presented in Section 3.3 below.

2.6 Sensitivity analysis of fitting procedure of bivariate distributions

The present study is based on a large data set of descriptors computed from experimentally measured image data. However, the question arises, how sensitive the presented procedure is to the amount of available data, and consequently, how many measurements are necessary in order to achieve a reasonable quality of fit. To answer these questions, a bootstrap sampling-based sensitivity analysis of the presented modeling approach is deployed [37]. This involves the quantitative analysis of model fits that are achieved on a data set containing only a fraction of the measured data, allowing for the analysis of the added value of an increasing amount of available data.

To analyze the sensitivity of the fit of the probability density $f_{d,s}$ that is based on some data $\mathcal{Y}$ (e.g., $\mathcal{Y} = D_{t,(\mathrm{d,\ s})}$ for some $t \in T$), first, a bootstrap sample $\widetilde{\mathcal{Y}}$ of size $n_\mathrm{b} \in \mathbb{N}$ is constructed by drawing $n_\mathrm{b} > 0$ data points uniformly at random from the

set $\mathcal{Y}$. Then, a second probability density $\widetilde{f}_{d,s}$ is fitted with the data in $\widetilde{\mathcal{Y}}$. In this manner, we can investigate the discrepancy (see below for further details) of $f_{d,s}$ to a fit $\widetilde{f}_{d,s}$ that has been achieved with fewer data.

In the present paper, the discrepancy between three probability densities $\widetilde{f}_{d,s}$ and $f_{d,s}$ is quantified in two ways. First, the absolute percentage errors APE_d and APE_s of the expected values of the marginal distributions are considered this is given by

$$\mathrm{APE}_d(f_{d,s}, \widetilde{f}_{d,s}) = \frac{|\int_0^\infty x(\widetilde{f}_d(x) - f_d(x))\,dx|}{|\int_0^\infty x f_d(x)\,dx|}, \tag{19}$$

$$\mathrm{APE}_s(f_{d,s}, \widetilde{f}_{d,s}) = \frac{|\int_0^\infty x(\widetilde{f}_s(x) - f_s(x))\,dx|}{|\int_0^\infty x f_s(x)\,dx|}, \tag{20}$$

where $f_d, f_s \colon \mathbb{R} \to [0,\infty)$ and $\widetilde{f}_d, \widetilde{f}_s \colon \mathbb{R} \to [0,\infty)$ are the marginal probability densities of $f_{d,s}$ and $\widetilde{f}_{d,s}$, respectively. In order to quantify not only the discrepancy of the marginal distributions, but also the discrepancy of the dependency structures of the marginal distributions, a second measure $L(f_{d,s}, \widetilde{f}_{d,s}) \in [0,2]$, is utilized. This measure compares the copula densities $c, \widetilde{c} \colon [0,1]^2 \to [0,\infty)$, of $f_{d,s}, \widetilde{f}_{d,s}$ by means of the L_1-norm and is given by

$$L(f_{d,s}, \widetilde{f}_{d,s}) = \int_0^1 \int_0^1 |c(x,y) - \widetilde{c}(x,y)|\,dx\,dy. \tag{21}$$

A value of $L(f_{d,s}, \widetilde{f}_{d,s})$ close to zero, corresponds to a high similarity, whereas a value close to two, corresponds to extreme dissimilarity.

By means of the outlined bootstrapping approach, we can investigate the goodness of fit in dependence of the number n_b of sampled data points. In other words, this approach enables us to assess the number of objects and, consequently, the number of measurements necessary to achieve the desired precision in our model fits; see Section 3.4.

3 Results and discussion

3.1 Segmentation

To evaluate the quality of the segmentation procedure described in Section 2.2, a combination of visual and quantitative analyses was performed. To do so, for five of the Camsizer images, a ground truth phase-wise segmentation $p^* \colon \mathcal{X} \to \{0,1,2\}$ was generated by using a much slower state-of-the-art segmentation modelfrom the field of machine learning [38] . However, this method is not feasible for inline segmentation due to its computational complexity in memory and time. The difference of this ground truth segmentation p^*, and the segmentation achieved with the method described in Section 2.2 is visualized in Figure 4. It can be observed that all objects are detected correctly, and differences in segmentations are due to small variations of the objects' outlines.

Furthermore, for a quantitative evaluation of the segmentation quality, the segmentations p from Section 2.2 are compared with p^* by means of the intersection over union (IoU) metric [39]. The IoU is defined as

$$\mathrm{IoU}(p, p^*) = \frac{|\{x \in \mathcal{X} \colon p(x) = 1 \text{ and } p^*(x) = 1\}|}{|\{x \in \mathcal{X} \colon p(x) = 1 \text{ or } p^*(x) = 1\}|}. \tag{22}$$

The segmentation method described in Section 2.2 achieves an average IoU score of 0.93 compared to the reference ground truth segmentation, indicating a high degree of agreement.

3.2 Classification

For object type classification, a random forest is trained as described in Section 2.4 on basis of 1854 descriptor vectors and corresponding hand labeled object classes. Note that this training data consists of descriptor vectors computed from image data derived from both Experiment A and Experiment E as well as all time steps in T. The grid-search to tune the hyperparameters leads to a random forest with 100 decision trees each of which have a maximal depth of 5. Moreover, $|J| = 5$ randomly chosen descriptors are considered per tree. The prediction quality of the random forest classifier that achieved the best results, is evaluated based on a second set of hand-labeled particle descriptors, again containing descriptors from all experiments and time steps, but not used in the training of the random forest. Specifically, the prediction quality of the classifier was evaluated based on descriptor vectors of 276 primary particles 133 chain-like agglomerates and 148 raspberry-like agglomerates. The corresponding confusion matrix is shown in Table 5 (left).

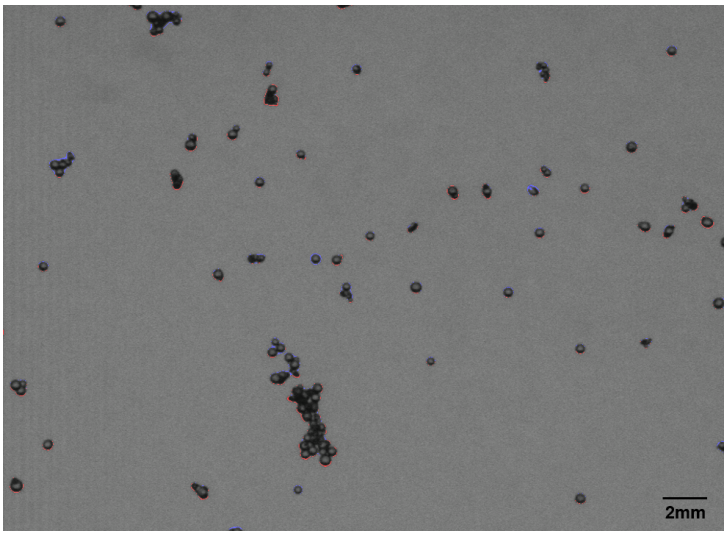

Figure 4: **Segmentation quality.** The difference between the reference ground truth segmentation p^* obtained using SAM [38] and the method described in Section 2.2 is shown. Pixels segmented by the proposed method but not by SAM are highlighted in red, while those segmented by SAM but not by the proposed method are shown in blue.

To assess the added value of utilizing a larger number of particle descriptors in the random forest classification compared to using only a few, a second, more interpretable classification approach based on only two particle descriptors was implemented. Specifically, the diameter d and eccentricity e were selected for this classification, as the diameter is a straight-forward descriptor to distinguish between primary particles and agglomerates. Furthermore, chain-like agglomerates tend to be much more elongated than raspberry-like agglomerates, thus, typically showing higher values of e. Based on the data used for the training of the random forest classifier, an optimal threshold for the area-equivalent diameter d was computed to distinguish primary particles from agglomerates. Note that it is enough to restrict the search for such a threshold to the observed values of d in the training data, ensuring both efficiency and straightforward computation. Subsequently, a second optimal threshold for the eccentricity descriptor e was determined to classify non-primary particles into chain-like and raspberry-like agglomerates. The resulting confusion matrix for this reference classifier, computed on the validation data, is shown in Table 5 (right). The random forest classification, operating on all 22 descriptors considered in the present paper, achieves a much better precision among all classes than the reference classifier which only considers the diameter d and eccentricity e. This justifies the use of a more complex, and thus, more time-consuming, classification procedure.

The influence of individual descriptors on the classification of particles/agglomerates can be measured by so-called Shapley values [40, 41, 42]. It turns out that for classifying primary particles, the most influential descriptor is roundness ψ, followed by the length of the major axis v_1 and the ratio of the radii of the inscribed and enclosing spheres r. For identifying chain-like agglomerates, roundness ψ remains the most influential descriptor, followed by the lengths of the minor axis v_2 and major axis v_1. Finally, for classifying raspberry-like agglomerates, the minor axis length v_2 is the most influential descriptor, followed by roundness ψ and again the major axis length v_1

Table 5: **Confusion matrices.** The confusion matrix for the random forest classifier (left) and of the corresponding reference classifier that is based on the diameter d and eccentricity e is shown. The term GT refers to the particle classes assigned by hand labeling.

		random forest		
		primary	chain	berry
GT	primary	275	0	0
	chain	0	123	10
	berry	0	11	137

		reference classifier		
		primary	chain	berry
GT	primary	265	0	11
	chain	13	96	24
	berry	0	37	111

3.3 Particle descriptor prediction

3.3.1 Temporal evolution of class sizes

In the following the agglomeration process in both Experiments A and E is modeled in terms of the size fractions of the three object classes over time. More specifically, for each experiment and each measured time step, the fraction of segmented foreground pixels belonging to the respective classes is first computed. Then, as described in Section 2.5.1, the regression

function from Equation (9) is fitted to model the evolution of class size fractions over time for each experiment individually. The resulting fitted regression functions are shown in Figure 5.

Focusing on primary particles in Experiments A and E, their number is decreasing as more and more primary particles being agglomerated, see Figure 5 (blue lines). Nevertheless, the fraction of primary particles does not vanish completely, due to two effects: 1) continuous feed of primary particles into the process, 2) intermediate breakage of formed agglomerates. Volume fractions of chain-like and raspberry-like agglomerates also attain steady values, although with different dynamics. Initially, a large number of chain-like agglomerates are formed (only few agglomeration events required) which are then integrated into larger, raspberry-like structures. In total, a slight prevalence of raspberry-like agglomerates over chain-like agglomerates is observed in Experiment A, that can be attributed to the higher mechanical stability of raspberry-like agglomerates (larger number of solid bridges with surrounding particles). Although the general trends are the same for Experiment E, it differs in its kinetics, i.e., agglomerate formation is slower in Experiment E than in Experiment A. This can be related to the operation conditions: Experiment E is operated at a higher gas inlet temperature and with a higher binder content. Consequently, the sprayed droplets will dry faster and form individual particles (called overspray) that do not contribute to the agglomeration, as these pre-dried droplets will not deposit on the primary particles or agglomerates.

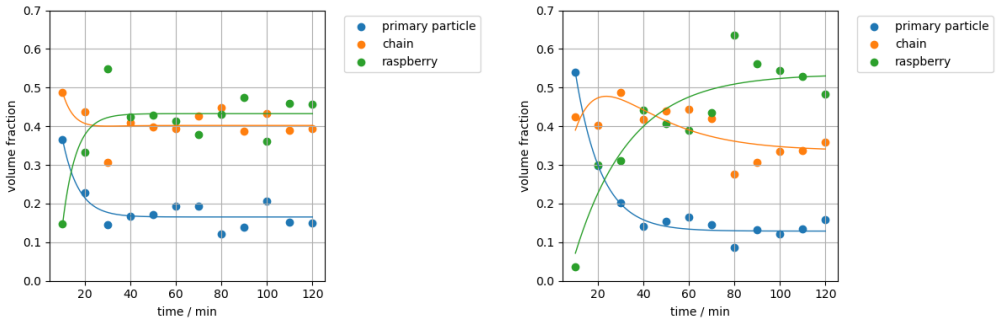

Figure 5: **Temporal evolution of classes sizes.** The temporal evolution of the area-weighted size fraction for different particle/agglomerate classes in Experiment A (left) and Experiment E (right) is shown. Circles represent the area-weighted fraction of particles/agglomerates observed in the image data. The fitted regression function (Equation (9)), derived in Section 3.3.1, is shown as curves in the corresponding colors.

3.3.2 Fitted univariate distributions over time

To analyze and model not only the volume fraction of these classes, but also the probability distributions of size and shape descriptors, Section 2.5.2 introduced a parametric modeling procedure for the univariate distributions of object size d and solidity s. We apply the described procedure to the data of each class and of Experiment A and E, to obtain the univariate fits $f_d^{\widehat{G}_d,\widehat{\omega}_{t,d}}$ and $f_s^{\widehat{G}_s,\widehat{\omega}_{t,s}}$ for each $t \in T$. The resulting best-fitting family and the corresponding parameters of the marginal distributions for each experiment, class and time step are presented in Table 6.

For the primary particles, no trend can be seen in the evolution of distribution parameters over time for both experiments. This indicates that the size and shape of these particles remain unchanged over time, which is expected for primary particles. Consequently, fitting a regression line provides no additional value as we assume that the univariate distributions are constant over time.

The time-dependent regression function given in Equation (9) is fitted to the distribution parameters given in Table 6, i.e., to distribution parameters of each experiment, each agglomerate type and each descriptor, by minimizing Equation (18). The resulting regression functions fitted to the distribution parameters in Table 6 for raspberry-like agglomerates are shown on the right side of Figure 6. The upper row shows the results corresponding to the area-equivalent diameter d, whereas the lower row shows the results corresponding to the solidity s. All parameters reach saturation after 30 minutes, with the exception of the parameters that describe the area-equivalent diameter of raspberry-like agglomerates in Experiment E. The fitted regression functions can be used to predict marginal distributions for unmeasured time steps.

Figures 6 a), b), e) and f) visualize the resulting marginal distributions of raspberry-like agglomerates for experiments A and E at the time steps 30 min and 120 min. Note that the parameters of the visualized probability densities (lines) have been obtained by using the prediction of the associated regression functions at these time steps. As can be observed, the distributions correspond well with the available data. Moreover, the regression functions have been used to predict distribution parameters for the probability distribution of d and s at the time step 75 min, i.e., for a time step for which no data is available. The corresponding probability densities are visualized in gray in Figure 6. Interestingly, the distributions obtained from the regression curves remain unchanged between these time steps, except for the area-equivalent diameter d of agglomerates in Experiment E. However, in Experiment E, the size of the agglomerates continues to increase after 30 min,

while their shape, i.e., their solidity s, remains constant. This behavior is to be expected based on the regression functions for the distribution parameters.

Table 6: **Fitted parametric univariate distributions.** Parametric families of univariate distributions that have been identified as suitable for modeling the univeriate probability densities of d and s for each class in Experiments A and E across all time steps, along with the fitted parameters for each time step (desc. = descriptor, dist. = distribution, para. = parameter).

desc.	dist.	par.	time step t /min											
			10	20	30	40	50	60	70	80	90	100	110	120
Experiment A; primary particle														
d	normal	$\mu =$	215.3	224.89	222.79	221.18	221.24	221.3	222.46	223.22	226.97	223.83	226.57	221.81
		$\sigma =$	29.01	28.53	29.6	29.58	27.9	27.77	27.82	27.24	29.6	30.34	29.08	29.05
s	normal	$\mu =$	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96
		$\sigma =$	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Experiment A; chain														
d	gamma	$\alpha =$	33.53	33.57	42.44	24.78	25.99	25.63	26.08	21.5	30.3	25.01	32.85	25.83
		$\beta =$	8.45	8.9	7.36	12.2	11.73	11.79	11.39	14.84	10.3	12.43	9.09	11.88
s	normal	$\mu =$	0.91	0.91	0.91	0.9	0.9	0.9	0.91	0.89	0.9	0.9	0.91	0.91
		$\sigma =$	0.04	0.04	0.06	0.05	0.05	0.06	0.06	0.07	0.05	0.06	0.06	0.06
Experiment A; raspberry														
d	log-normal	$\mu =$	0.11	0.2	0.21	0.2	0.22	0.22	0.19	0.19	0.21	0.21	0.17	0.18
		$\sigma =$	417.45	468.71	479.03	464.03	473.36	474.12	469.77	453.2	478.91	455.65	459.8	465.19
s	normal	$\mu =$	0.89	0.87	0.86	0.88	0.87	0.86	0.87	0.87	0.86	0.87	0.87	0.87
		$\sigma =$	0.04	0.07	0.07	0.06	0.07	0.08	0.06	0.05	0.06	0.08	0.06	0.06
Experiment E; primary particle														
d	normal	$\mu =$	208.44	213.41	217.9	219.66	217.17	219.35	216.03	217.14	217.75	217.13	215.05	213.89
		$\sigma =$	31.73	28.96	28.5	29.66	31.07	27.61	29.96	27.58	27.51	31.68	31.77	30.52
s	normal	$\mu =$	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96
		$\sigma =$	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.02
Experiment E; chain														
d	log-normal	$\mu =$	0.18	0.2	0.19	0.2	0.2	0.21	0.21	0.18	0.21	0.22	0.19	0.23
		$\sigma =$	256.34	280.19	293.77	303.7	304.58	310.27	300.08	306.56	292.69	307.77	305.23	295.43
s	normal	$\mu =$	0.91	0.91	0.9	0.9	0.9	0.89	0.9	0.9	0.9	0.89	0.89	0.9
		$\sigma =$	0.04	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.05	0.06
Experiment E; raspberry														
d	log-normal	$\mu =$	0.1	0.19	0.18	0.18	0.19	0.19	0.19	0.21	0.18	0.21	0.21	0.21
		$\sigma =$	369.2	467.29	463.34	473.1	449.76	449.77	464.17	481.65	473.71	497.88	482.48	478.38
s	normal	$\mu =$	0.91	0.85	0.87	0.86	0.87	0.87	0.87	0.86	0.87	0.86	0.86	0.85
		$\sigma =$	0.01	0.06	0.06	0.05	0.08	0.06	0.06	0.06	0.07	0.05	0.07	0.06

We obtain similar results for chain-like agglomerates. For the solidity s, the distribution parameters remain almost unchanged after 30 min for both experiments, see Table 6. The same applies for the area-equivalent diameter d and Experiment A. However, similar to the raspberry-like agglomerates, the area-equivalent diameter in Experiment E continues to increase beyond 30 minutes, as indicated by the rising value of σ in Table 6. Nevertheless, this increase is less pronounced than in the raspberry-like agglomerates and nearly reaches saturation after 60 minutes.

The general preservation of shape (solidity) with respect to the operation conditions is not surprising: Due to fluidization, agglomerates and recently attached primary particles undergo constant collisions with other agglomerates and the apparatus walls. The mechanical stress is sufficient to overcome the cohesive forces required to break single bridges. This favors agglomerate structures with low surface areas and large numbers of contact points between primary particles, i.e., the observed raspberry-like shape. Although the shape is mostly unchanged, the dynamics of agglomerate formation, agglomerate size and the intermediate volume fraction differ significantly, compare Figure 5.

3.3.3 Fitted bivariate distributions over time

So far, only univariate distributions of d and s have been considered, but these do not model the dependency of the two descriptors. Thus, we want first to investigate the dependency of area-equivalent dimeter and solidity for all three classes observed in Experiments A and E. A non-parametric measure for dependence is empirical Kendall's tau, see [22], which takes values in the interval $[-1, 1]$, where a value close to zero indicates that there is no clear positive or negative correlation between the area-equivalent diameter and solidity.

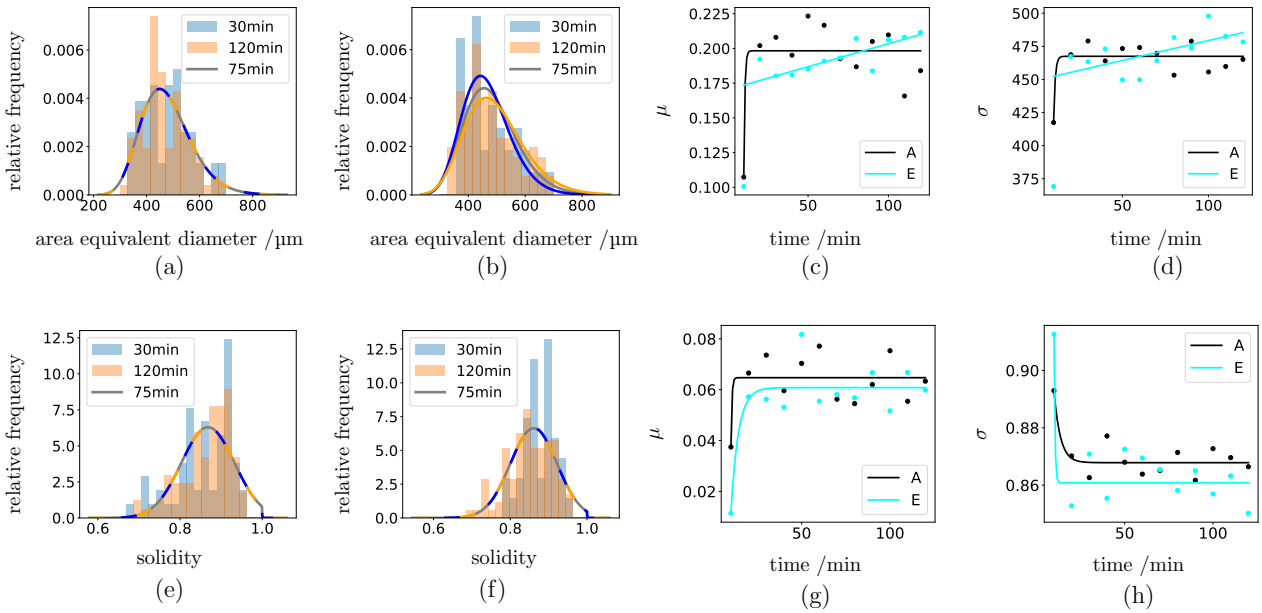

Figure 6: **Temporal evolution of marginal distributions.** Parametric modeling of marginal distributions. The marginal distributions of the area-equivalent diameter d (upper row) and solidity s (lower row) for both, Experiment A (column 1) and Experiment E (column 2) are shown for the time steps 30 min, 75 min and 120 min. The distributions of all time steps are obtained by parameter regression. The parameters of the marginal distributions and a fitted regression line are shown (column 3-4) for both experiments. The regression makes it possible to make statements about unobserved time steps, e.g. for 75 minutes for which no data is available.

The empirical Kendall's tau values between d and s for primary particles, averaged over all time steps, is 0.10 across both experiments. For chain-like agglomerates, the corresponding average value of empirical Kendall's tau over all time steps are -0.35 and -0.38 for Experiments A and E, respectively. Analogously, we determined the averages of -0.47 and -0.40 for raspberry-like agglomerates for Experiments A and E over all time steps, respectively. Due to the average Kendall's tau being close to 0, we assume independence of d and s for primary particles.

Recall that copulas can be used to model the dependencies between two descriptors. Section 2.5.2 outlined how the marginal distributions and an Archimedean copula can be used to model the bivariate distribution of d and s. However, since we assume independence for the two descriptors in the case of primary particles, we solely employ copulas to model the bivariate distributions of d and s for agglomerates.

To determine the best-fitting copula family and the corresponding parameter for each time step of the agglomerates, we apply the methods from Section 2.5.2 to the data of chain-like and raspberry-like agglomerates for Experiment A and Experiment E to obtain a sequence of parametric copulas $c_{\mathrm{d,s}}^{\widehat{Z},\widehat{\theta}_t}$, see Table 7.

Table 7: **Fitted parametric copula functions.** Archimedean copulas that have been identified as suitable for modeling the bivariate probability of d and s for raspberry-like and chain -like agglomerates in Experiment A and Experiment E across all time steps, along with the fitted parameters for each time step (rot. = rotation, exp. =experiment, par. = parameter, ali. = Ali-Mikhail-Haq).

exp.	type	copula	rot.	par.	time step t /min											
					10	20	30	40	50	60	70	80	90	100	110	120
A	chain	ali	90	$\theta =$	0.87	0.87	0.98	1.0	0.98	0.99	0.99	0.99	1.0	1.0	0.99	0.96
A	raspberry	clayton	270	$\theta =$	0.78	1.88	1.72	1.55	1.83	1.79	1.53	1.97	2.2	1.92	1.53	1.8
E	chain	ali	90	$\theta =$	0.79	0.96	0.98	1.0	0.99	0.99	0.98	0.97	0.96	0.99	0.96	0.98
E	raspberry	clayton	270	$\theta =$	0.16	1.22	1.5	1.4	1.77	1.81	1.6	1.17	1.58	1.55	1.84	1.81

Figure 7 visualizes exemplarily chosen bivariate probability densities of d and s for primary particles, chain-like agglomerates and raspberry-like agglomerates. A visual inspection of the densities for each class shows that a linear combination of the three densities represents the distribution of all scatter points shown in Figure 7 (first column) well.

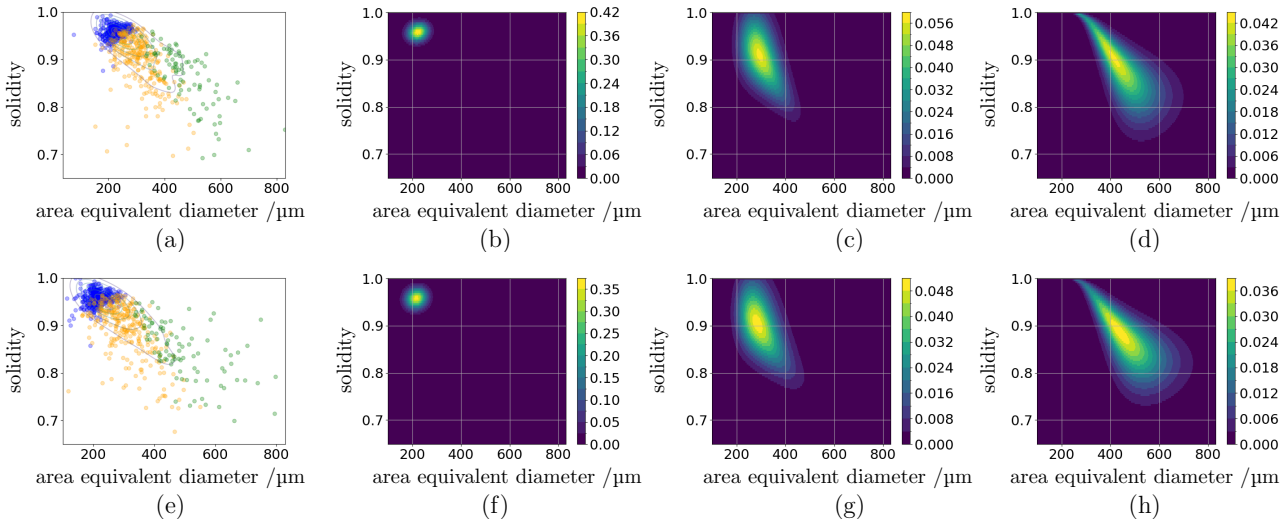

Figure 7: **Bivariate distribution of descriptors.** Bivariate distribution of area-equivalent diameter and solidity for particles/agglomerates observed in Experiments A (upper row) and E (lower row). The bivariate distributions of all measured particles is represented using iso-lines of the probability density obtained by means of a kernel density estimation (column 1) at time step 120 min. Furthermore, a decomposition of this distribution into the parametric distributions of the classes of primary particles, chain-like agglomerates, and raspberry-like agglomerates is shown (columns 2-4).

As for the distribution parameters of the marginal distributions of d and s, temporal changes of copula parameters can be analyzed for agglomerates. Therefore, time-dependent regression functions as given in Equation (9) are fitted to the values of distribution parameter by minimizing Equation (18). The resulting regression functions fitted to the copula parameters in Table 7 for raspberry-like agglomerates observed in Experiments A and E are shown in Figures 8d) and h). The parameters of Experiment A reach saturation after 20 min, while the parameters of Experiment E reach saturation after 40 min. In addition, the bivariate probability densities of d and s for raspberry-like agglomerates after 30 and 120 min is shown in the first and third column, where the copula parameters of the distributions have been determined by means of regression. Once again, no significant changes can be observed in the distributions for Experiment A. However, in Experiment E, an even more pronounced shift in the distributions can be seen, which is mainly caused by the parameter evolution of the area-equivalent diameter, see Section 3.3.2.

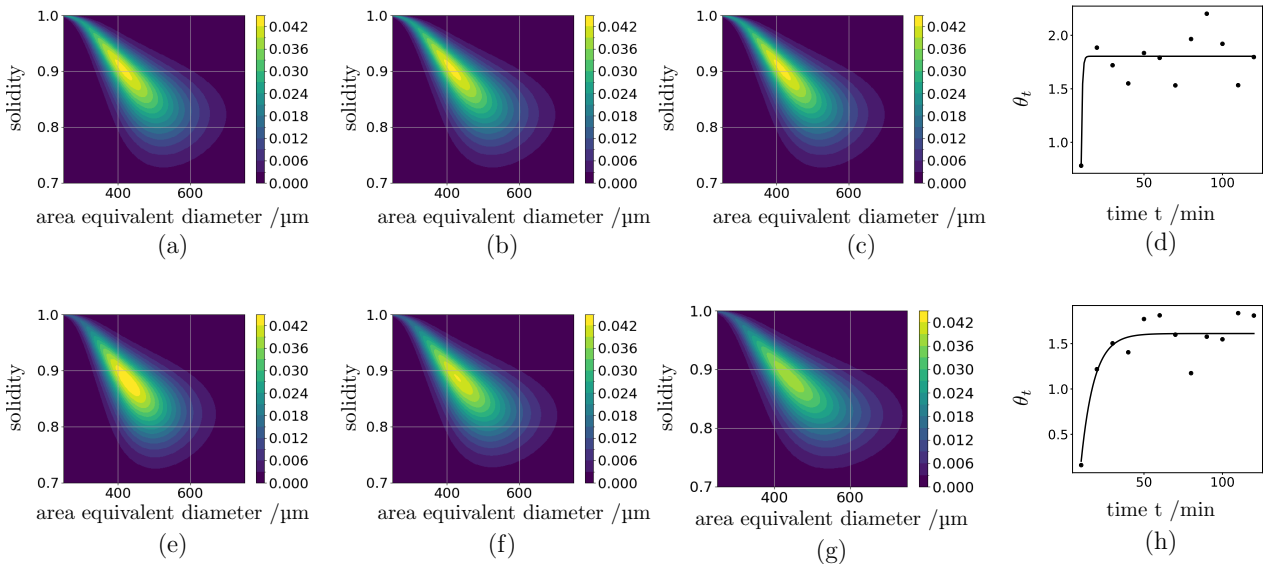

Figure 8: **Temporal evolution of marginal distributions.** Evolution of bivariate probability density of area-equivalent diameter and solidity in time for Experiment A (upper row) and E (lower row). The bivariate probability density for raspberry-like agglomerates is shown for the time steps of 30 min (column 1) and 120 min (column 3). Furthermore, the predicted bivariate probability density for raspberry-like agglomerates is shown for the time step of 75 min (column 2) for which no measured data is available.

Similarly to the univariate case, the fitted regression functions can be used to estimate copula parameters and consequently bivariate distributions of descriptor vectors for unmeasured time steps. The predicted bivariate distributions for raspberry-like agglomerates are shown in Figures 8b) and f) for Experiment A and Experiment E at 75 min, respectively.

For chain-like agglomerates, the parameters in both experiments reach saturation after 30 min. Results therefore indicate that with respect to structure formation, dynamic equilibrium between agglomeration of primary particles into chain-like agglomerates, transition from chains-like agglomerates to raspberry-like agglomerates and the breakage of too large raspberry-like agglomerates is reached. This timescale is also in line with results on average agglomerate size presented in [25]. The previously presented results on modeling the evolution of particle and agglomerate structure distributions over time provide valuable insights into the process.

3.4 Sensitivity of fitting procedure of copula-based bivariate distributions

As the previous results were obtained using a large amount of data that is not available in an online or real-time measurement situation, the question arises how sensitive these results are with respect to the number of evaluated objects. Or: How many objects must be evaluated before a reliable and robust estimate of the multi-dimensional structure of agglomerates is obtained?

In order to investigate the sensitivity of the model fitting procedure described in Section 3.3, we deployed the sensitivity analysis procedure described in Section 2.6 to the particle descriptors measured in Experiment E at time 120 min. More precisely, for each object class we conducted the sensitivity analysis described in Section 2.6 by setting $\mathcal{Y}$ as the set of all descriptor vectors $(d, s) \in [0, \infty) \times [0, 1]$ of particles/agglomerates of the respective class, measured in Experiment E at time 120 min. Then, for each $n_\mathrm{b} \in \{5, 20, 35, \ldots, 140\}$ a total of 1 000 bootstrap samples $\widetilde{\mathcal{Y}}$ of size n_b are independently generated from $\mathcal{Y}$ and the corresponding models $\widetilde{f}_{d,s}$ are fitted. Note that since we assume that the parametric families for the marginal distributions and for the copula do not change for different time steps within one experiment and object class, the parametric families associated with $\widetilde{f}_{d,s}$ are chosen in accordance with Table 3 and Table 7. More precisely, when fitting $\widetilde{f}_{d,s}$ to $\widetilde{\mathcal{Y}}$, we skip the search for optimal parametric families for both the marginal distributions and the copula, as we take the families from $f_{\mathrm{d,s}}$ given in Table 3 and Table 7. We then optimize only their parameters, $\omega_d, \omega_s \in \mathbb{R}^2$ and $\theta \in \Theta$ with respect to $\widetilde{\mathcal{Y}}$, using maximum likelihood estimation [36].

The results of this analysis are presented in Figure 9. On the left side and in the middle, the similarity of $f_{d,s}$ and $\widetilde{f}_{d,s}$ are shown by means of the APE_d and APE_s introduced in Equation (19) and Equation (20). It can be observed that accurate modeling of primary particles requires only a few observations. In contrast, modeling chain-like agglomerates necessitates a larger number of observations, while the modeling of raspberry-like agglomerates requires the most observations. This aligns with the observation that raspberry-like structures are much more complex than primary particles or chain-like agglomerates. On the right side of Figure 9, we show the sensitivity of the procedure to fit the dependency structure of the probability density in dependence of the amount of available data. Since the dependency structure of d and s for primary particles is assumed to be independent, they are excluded from this analysis. As before, the dependency structure of d and s for raspberry-like agglomerates is more sensitive than that of chain-like agglomerates.

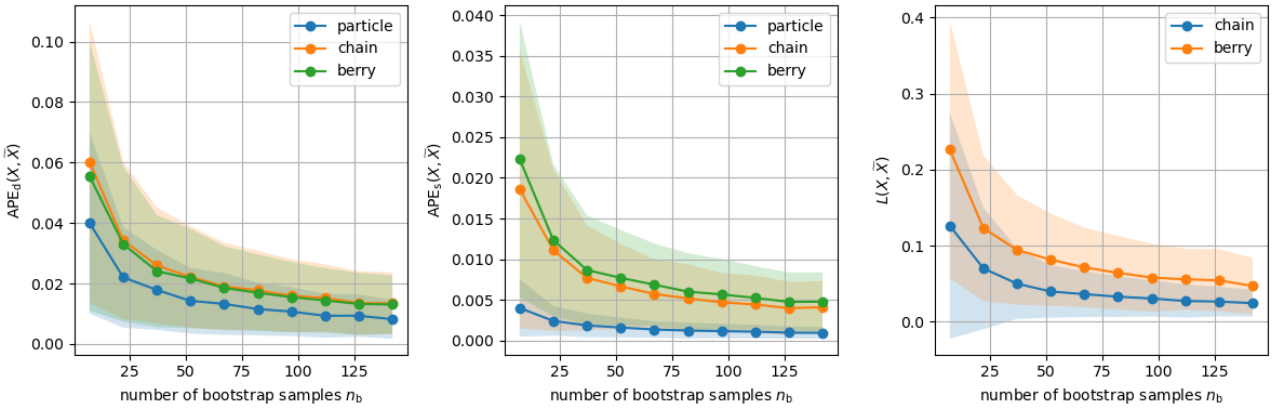

Figure 9: **Bootstrap sampling-based sensitivity.** For all particle types, the similarity of $f_{d,s}$ computed by means of all data descriptor vectors measured in Experiment E at time step 120 min and $\widetilde{f}_{d,s}$ computed by means of a subset of n_b descriptor vectors is shown. On the left and in the middle, the absolute percentage loss (see Equation (19) and Equation (19)) is shown for different sizes n_b. On the right-hand side, the similarity of $f_{d,s}$ and $\widetilde{f}_{d,s}$ is shown by means of the integral over the absolute difference of the respective copula functions, see Equation (21). All experiments are conducted, 1 000 times, the solid lines show the arithmetic mean, while the shaded area shows the standard deviation.

As raspberry-like agglomerates have the most complex structure and appear less frequently in our data, estimates on the required number of identified objects are expressed with respect to this agglomerate class. Given a desired expected error, in terms of APE_d and APE_s, the minimum number of detected agglomerates can be predicted. Figure 9 (Experiment E) shows that approximately 70 raspberry-like agglomerates need to be measured in order to achieve in expectation an APE_d and APE_s of less than 2 %.

For Experiment E with an average count of about 3.5 raspberry-like agglomerates per image, this corresponds to about 20 images. Given typical acquisition rates (frame rates) of about 60 images per second, this means that after approximately a third of a second of measurement time the quantity of information necessary for such a precision can be acquired.

These results are exceptionally important for implementing online measurement of agglomerate structure formation and online model-learning and process adaptation, e.g., closed-loop (feedback) control of fluidized bed spray agglomeration processes for defined agglomerate structures. In particular, it enables almost immediate action to steer synthesized agglomerate structures into preferable directions; furthermore, deviations and disturbances can be quickly identified and severe malfunctions (e.g., break-down of fluidization due to excessive agglomeration) can be prevented.

4 Conclusion

In this paper, the agglomeration process of glass beads in a SFB agglomeration over time is quantitatively investigated and modeled. In particular, the structure formation is investigated by analyzing inline images sequences followed by time-dependent multivariate stochastic modeling of size and shape descriptors of particles/agglomerates.

Particles and agglomerates are automatically segmented from image data, from which various geometrical and textural descriptors are computed. These descriptors are used to classify segmented objects as primary particles, chain-like agglomerates and raspberry-like agglomerates. For this task, a fast and robust random forest classifier is successfully applied. Subsequently, parametric bivariate distributions are fitted for each class and time step. This allows the regression of model parameters, which enables us to make statements about the distribution of descriptor vectors at unobserved time steps and predict the time-dependent multivariate descriptor distribution of these particle classes. In addition, the volume fraction of each class over time is investigated. The results show that no further agglomeration can be observed after 30 minutes with a small amount of binder. With a higher amount of binder, however, further agglomeration is observed after 30 minutes.

In addition, a sensitivity analysis is performed to quantify the amount of data required in order to adequately model the bivariate distributions that characterize the state of agglomeration. It is shown that a higher amount of data is required for raspberry-like agglomerates due to their more complex shape and less-frequent occurrence than chain-like agglomerates and primary particles. The results of the sensitivity analysis show the minimum data required for online tracking of agglomerate formation dynamics. This opens the door to online monitoring of structure formation and feedback control of SFB agglomeration with respect to disturbance identification and rejection. As a result, stable operation can be maintained while ensuring predefined product properties such as the re-hydration capacity and kinetics of the agglomerated material.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the authors.

Acknowledgements

Funding of this work by Deutsche Forschungsgemeinschaft (DFG) (project IDs 504524147 and 504580586) within the priority programme PP 2364 Autonomous Particle Processes is gratefully acknowledged.

References

[1] D. G. Bika, M. Gentzler, and J. N. Michaels. “Mechanical properties of agglomerates”. In: *Powder Technology* 117 (2001), pp. 98–112.

[2] T. Gluba. “Drum granulation conditions for raw material with different particle size distributions”. In: *Handbook of Powder Technology* 10 (2001), p. 717.

[3] W. Pietsch. *Agglomeration Processes*. WILEY-VCH, 2002.

[4] C. Turchiuli, Z. Eloualia, N. E. Mansouri, and E. Dumoulin. “Fluidised bed agglomeration: Agglomerates shape and end-use properties”. In: *Powder Technology* 157 (2005), pp. 168–175.

[5] A. Kataria. “An analysis of drug migration during drying of granules as an underlying cause of content non-uniformity in wet granulation”. MA thesis. Rutgers, The State University of New Jersey, 2018.

[6] B. Liu, J. Wang, J. Zeng, L. Zhao, Y. Wang, Y. Feng, and R. Du. “A review of high shear wet granulation for better process understanding, control and product development”. In: *Powder Technology* 381 (2021), pp. 204–223.

[7] F. Robin, H. P. Schuchmann, and S. Palzer. “Dietary fiber in extruded cereals: Limitations and opportunities”. In: *Trends in Food Science & Technology* 28 (2012), pp. 23–32.

[8] Z. Wang, Z. Pan, D. He, J. Shi, S. Sun, and Y. Hou. “Simulation modeling and of a pharmaceutical and tablet manufacturing and process via wet and granulation”. In: *Complexity* 2019.1 (2019), p. 3659309.

[9] L. Fries, S. Antonyuk, S. Heinrich, G. Niederreiter, and S. Palzer. “Product design based on discrete particle modeling of a fluidized bed granulator”. In: *Particuology* 12 (2014), pp. 13–24.

[10] S. Palzer. “Agglomeration of pharmaceutical, detergent, chemical and food powders — Similarities and differences of materials and processes”. In: *Power Technology* 206 (2011), pp. 2–17.

[11] H. Schuchmann. “Production of instant foods by jet agglomeration”. In: *Food Control* 6 (1995), pp. 95–100.

[12] N. Jinapong, M. Suphantharika, and P. Jamnong. “Production of instant soymilk powders by ultrafiltration, spray drying and fluidized bed agglomeration”. In: *Journal of Food Engineering* 84 (2008), pp. 194–205.

[13] E. Otto, R. Dürr, A. Kienle, A. Bück, and E. Tsotsas. “Dynamic modeling of particle size and porosity distribution in fluidized bed spray agglomeration”. In: *Computer Aided Chemical Engineering* 53 (2024), pp. 163–168.

[14] E. Otto, A. Ajalova, A. Bück, E. Tsotsas, and A. Kienle. “Population balance modeling of particle size and porosity in fluidized bed spray agglomeration”. In: *Industrial & Engineering Chemistry Research* 63 (2024), pp. 17545–17556.

[15] K. Chen, Z. Li, S. Akbas, and E. Tsotsas. “Monte Carlo modeling of particle agglomeration during polymer pyrolysis in bubbling fluidized bed”. In: *Fuel* 367 (2024), p. 131487.

[16] X. Deng, Z. Huang, W. Wang, and R. N. Davé. “Investigation of nanoparticle agglomerates properties using Monte Carlo simulations”. In: *Advanced Powder Technology* 27 (2016), pp. 1971–1979.

[17] S. N. Rogak, R. C. Flagen, and H. V. Nguyen. “The mobility and structure of aerosol agglomerates”. In: *Aerosol Science and Technology* 18 (1993), pp. 25–47.

[18] A. Buades, B. Coll, and J.-M. Morel. “Non-local means denoising”. In: *Image Processing On Line* 1 (2011), pp. 208–212.

[19] N. Otsu. “A threshold selection method from gray-level histograms”. In: *Automatica* 11.285-296 (1975), pp. 23–27.

[20] T. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, 2009.

[21] R. B. Nelsen. *An Introduction to Copulas*. Springer, 2006.

[22] H. Joe. *Dependence Modeling with Copulas*. CRC Press, 2014.

[23] O. Furat, T. Leißner, K. Bachmann, J. Gutzmer, U. Peuker, and V. Schmidt. “Stochastic modeling of multidimensional particle properties using parametric copulas”. In: *Microscopy and Microanalysis* 25.3 (2019), pp. 720–734.

[24] M. Weber, T. Wilhelm, and V. Schmidt. “Multidimensional characterization of time-dependent image data: A case study for the peripheral nervous system in aging mice”. In: *Image Analysis & Stereology* 40.2 (2021), pp. 85–94.

[25] A. Ajalova, T. Hoffmann, and E. Tsotsas. “Correlation between Process parameters and Structure-Property of Agglomerates Produced in a Spray Fluidized bed”. In: *Proceedings of the 23rd International Drying Symposium (Wuxi (CHN))*. 2024.

[26] S. N. Chiu, D. Stoyan, W. S. Kendall, and J. Mecke. *Stochastic Geometry and its Applications*. 3rd. J. Wiley & Sons, 2013.

[27] W. K. Pratt. *Introduction to Digital Image Processing*. CRC Press, 2013.

[28] L. Kenna. “Eccentricity in ellipses”. In: *Mathematics Magazine* 32.3 (1959), pp. 133–135.

[29] T. Allen. *Powder Sampling and Particle Size Determination*. Elsevier, 2003.

[30] H. G. Merkus. *Particle Size Measurements: Fundamentals, Practice, Quality*. Springer, 2009.

[31] K. Benkrid, D. Crookes, and A. Benkrid. “Design and FPGA implementation of a perimeter estimator”. In: *Proceedings of the Irish Machine Vision and Image Processing Conference*. Ed. by P. J. Morrow and B. W. Scotney. Vol. 51. 2000.

[32] R. Wang, A. K. Singh, S. R. Kolan, and E. Tsotsas. “Fractal analysis of aggregates: Correlation between the 2D and 3D box-counting fractal dimension and power law fractal dimension”. In: *Chaos, Solitons & Fractals* 160 (2022), p. 112246.

[33] G. Landini. “Fractals in microscopy”. In: *Journal of Microscopy* 241.1 (2011), pp. 1–8.

[34] C. Zhang and Y. Ma. *Ensemble Machine Learning*. Vol. 144. Springer, 2012.

[35] N. L. Johnson, S. Kotz, and N. Balakrishnan. *Continuous Univariate Distributions, Volume 1*. J. Wiley & Sons, 1994.

[36] M. Aitkin. *Statistical Inference: an Integrated Bayesian/likelihood Approach*. CRC Press, 2010.

[37] P. I. Good. *Resampling Methods*. Springer, 2006.

[38] A. Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A. C. Berg, W.-Y. Lo, A. C. Berg, W.-Y. Lo, D. Piotr, and R. Girshick. “Segment anything”. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2023, pp. 4015–4026.

[39] P. Jaccard. “The distribution of the flora in the alpine zone. 1”. In: *New Phytologist* 11.2 (1912), pp. 37–50.

[40] S. M. Lundberg and S.-I. Lee. “s”. In: *Advances in Neural Information Processing Systems*. Ed. by I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett. Vol. 30. Curran Associates, Inc., 2017.

[41] S. M. Lundberg, G. Erion, H. Chen, A. DeGrave, J. M. Prutkin, B. Nair, R. Katz, J. Himmelfarb, N. Bansal, and S.-I. Lee. “From local explanations to global understanding with explainable AI for trees”. In: *Nature Machine Intelligence* 2.1 (2020), pp. 2522–5839.

[42] Z. Zhang, Y. Dai, P. Xue, X. Bao, X. Bai, S. Qiao, Y. Gao, X. Guo, Y. Xue, Q. Dai, B. Xu, and L. Kang. "Prediction of microvascular obstruction from angio-based microvascular resistance and available clinical data in percutaneous coronary intervention: an explainable machine learning model". In: *Scientific Reports* 15.1 (2025), p. 3045.