Stochastic Modeling of Multidimensional Particle Properties Using Parametric Copulas

Orkun Furat\textsuperscript{1,*}, Thomas Leißner\textsuperscript{2}, Kai Bachmann\textsuperscript{3}, Jens Gutzmer\textsuperscript{3}, Urs Peuker\textsuperscript{2}, Volker Schmidt\textsuperscript{1}

\textsuperscript{1}Institute of Stochastics, Ulm University, D-89069 Ulm, Germany
\textsuperscript{2}Institute of Mechanical Process Engineering and Mineral Processing, Technische Universität Bergakademie Freiberg, D-09599 Freiberg, Germany
\textsuperscript{3}Helmholtz Institute Freiberg for Resource Technology, Helmholtz-Zentrum Dresden-Rossendorf, D-01328 Dresden, Germany

Abstract

In this paper, a prediction model is proposed which allows the mineralogical characterization of particle systems observed by X-ray micro tomography (XMT). The model is calibrated using 2D image data obtained by a combination of scanning electron microscopy (SEM) and energy dispersive X-ray spectroscopy (EDS) in a planar cross-section of the XMT data. To reliably distinguish between different minerals the model is based on multidimensional distributions of certain particle characteristics describing, e.g., their size, shape and texture. These multidimensional distributions are modeled using parametric Archimedean copulas, since other approaches like kernel density estimation require much larger sample sizes and are thus less practical. Parametric copulas have the advantage of describing the correlation structure of complex multidimensional distributions with only a few parameters. With the help of such distributions the proposed prediction model is able to distinguish between different types of particles among the entire XMT image.

Keywords and Phrases: X-ray micro tomography (XMT), mineral liberation analyzer (MLA), stereology, multidimensional particle characterization, parametric copula

1 Introduction

Processes which separate mixtures of particles based on criteria like size, shape or chemical composition of particles are used in many applications. Often the separation quality can have a critical influence on subsequent processing steps. For example, in the mining industry an essential step of ore dressing deals with separation processes which remove...
unwanted minerals from a system of particles with sizes smaller than 1 mm such that minerals of interest remain. In order to evaluate the separation success a mineralogical characterization of the particles prior and after separation is necessary. Common methods to characterize particles with respect to size, shape and composition rely on two-dimensional (2D) techniques describing a three-dimensional (3D) reality. For example, a combination of scanning electron microscopy (SEM) and energy dispersive X-ray spectroscopy (EDS) can be used (Sunderland and Gottlieb, 1991) to achieve a mineralogical characterization of ore samples. However, SEM-EDS is limited to 2D profiles of samples, which makes the characterization of a whole 3D sample rather expensive or even impossible - at least in a non-destructive manner. Another approach, namely X-ray micro tomography (XMT) provides 3D image data depicting the morphology of particles in a sample. This allows a quantitative analysis of particle systems without the stereological bias of 2D measurement techniques (Wang et al., 2017; Reyes et al., 2018). Recently, Su and Yan (2018) characterized sand particles observed in XMT with various shape descriptors and utilized spherical harmonic functions for both a parametric representation of single particles and a basis for stochastic modeling, see also Feinauer et al. (2015). Besides the morphology of particles, XMT provides information about local material specific constants. More precisely, the grayscale values of XMT images are related to the mass density of the observed material. When the considered minerals have distinct mass densities, the contrast-information from XMT can be utilized to distinguish between them. However, Furat et al. (2018) showed that grayscale values suffice only in a limited way to characterize the observed material in XMT data.

Therefore, for achieving a 3D mineralogical characterization of a sample it is necessary to utilize more information from the XMT data than solely the grayscale values. If we know, for example, that particles composed from a certain type of mineral are more spherical than others, we can utilize the additional information about their shape, which can be obtained from XMT as well, to characterize them.

In the present paper, we consider a sample which mainly comprises of the minerals zinnwaldite and quartz, though these particles can have imperfections consisting of topaz, muscovite, kaolinite and others, see Leißner et al. (2016). We propose a prediction model which can characterize 3D particles from XMT based on their size, shape and grayscale values. The calibration of this prediction model requires SEM-EDS data from only one planar cross-section of the sample. In order to be able to compute size and shape characteristics of XMT image data a particle-wise segmentation of the image data is required, such that it is possible to identify single particles in the XMT image. First, we give a short overview of the image processing steps which were necessary to obtain such a segmentation of our image data. Furthermore, we present several characteristics which can describe the size, shape and grayscale texture of particles. The mineralogical characterization of 3D particles from XMT images will be done based on these characteristics. Additionally to the XMT image, we have SEM-EDS data of the same sample, which lies in a planar cross-section of the XMT image. By utilizing the mineralogical characterization of SEM-EDS we know the mineralogical composition of the 3D particles that hit this cross-section. Therefore, we are able to link the size, shape and grayscale characteristics of these 3D particles to different minerals. The prediction model, proposed in the present paper, requires the probability density functions of the considered particle characteristics, for each type of mineral. Since multiple characteristics are necessary to characterize particles, because only grayscale information of the particles does not suffice for characterization, we need multidimensional probability densities of the considered
characteristics. An easy way to determine probability densities are so-called kernel density estimators (Scott, 2015), since they do not require the search for a suitable parametric family of distributions (e.g. normal distribution, exponential distribution, etc.). However, due to the “curse of dimensionality”, kernel density estimators require huge sample sizes for determining multidimensional densities which makes them less practical. Therefore, we rigorously show how to determine multidimensional densities by fitting parametric families of distributions to the data. At first we fit one-dimensional parametric densities to single characteristics and, in a second step, we use so-called Archimedean copulas (Nelsen, 2006) to determine joint densities of the considered particle characteristics. With the help of the fitted multidimensional densities we propose a prediction model which misclassifies only 2.7% of zinnwaldite particles and 5.7% of quartz particles. In a further step, we validate the prediction model by comparing its predictions with the characterization of SEM-EDS data at a spatially different cross-section which was not used for calibrating the prediction model.

The rest of this paper is organized as follows. In Section 2.1, we briefly recall some techniques of image processing and segmentation which we recently used in Furat et al. (2018). Then, in Section 2.2, various particle characteristics are explained, which are stochastically modeled later on in Section 2.3. Therefore, in Section 2.3.2, particular emphasis is put on the explanation of the copula approach to parametric modeling of multidimensional probability densities, and, in Section 3.1.1, the copula-based classification model itself as well its validation is explained. Finally, Section 4 concludes.

2 Materials and Methods

2.1 Image processing

2.1.1 3D XMT data preprocessing and segmentation

In this section, we give a short overview of the image preprocessing steps and the particle-wise segmentation of the 3D XMT data considered in the present paper, for more details see Furat et al. (2018). The first step consists of the reduction of noise in the XMT data with a non-local means denoising algorithm, see Buades et al. (2005). In contrast to the commonly used Gaussian filter, this nonlinear image filter has the advantage of smoothing homogeneous regions while preserving edges. In order to separate the particles from the background we use the local adaptive Sauvola thresholding algorithm, which determines binarization thresholds for each voxel based on its local neighborhood, see Shafait et al. (2008). Such local thresholding techniques can produce better results than global thresholds, due to globally inconsistent grayscale values in XMT images. One crucial and nontrivial task in image processing is the particle-wise segmentation, which allows the extraction of single particles from the image data for quantitative analysis. In a first step, we use a marker-based watershed algorithm, see Spettl et al. (2015), to get an initial segmentation. This results into some particles being wrongly separated into multiple fragments. To remove such oversegmentations we train a neural network (Hastie et al., 2009) to decide whether adjacent segments should be merged or not. To make this decision the neural network is supplied with information regarding the local morphology and grayscale values around two adjacent fragments. By applying the neural network as a post-processing step on the initial segmentation, we receive our final segmentation, see Figure 1b.
2. MATERIALS AND METHODS

2.1 Modeling of Particle Properties

2.1.2 2D SEM-EDS data and registration

In addition to the 3D XMT image data, SEM-EDS data was considered by Furat et al. (2018). The latter was obtained with a mineral liberation analyzer (MLA), from the same sample at two spatially different planar sections. An illustration of the SEM-EDS data can be seen in Figure 2a. It provides 2D information about the morphology of the particles in planar sections, but in contrast to the XMT data, see Figure 2b, it also provides information about the mineralogical composition of particles. For example, in the false color image of Figure 2a the blue phase indicates zinnwaldite.

Thus, by localizing the 2D SEM-EDS data in the segmented 3D image (registration), see Figure 1a, we have knowledge about the mineralogical composition of each 3D particle that intersects with the planar SEM-EDS data. Due to the particle-wise segmentation described in Section 2.1.1 we also know the 3D morphology of each of these particles. Furthermore, we can link the particle-wise segmentation with the grayscale values of the 3D XMT image, since the grayscale values provide valuable information about the composition of particles, as show in Furat et al. (2018). Therefore, both the grayscale values and the morphology of particles in the XMT image that hit the SEM-EDS plane allow...
2. MATERIALS AND METHODS

Modeling of Particle Properties

us to adjust a classifier that can predict the mineralogical composition of an arbitrary particle solely based on information gained from the XMT image.

To be precise, let $I: W \subset \mathbb{Z}^3 \to [0, 1]$ be the grayscale XMT image, where $W$ is a cuboid observation window. The 2D SEM-EDS data can be regarded as a map $L: H \cap W \subset \mathbb{Z}^3 \to \{0, 1, 2, \ldots\}$, where $H$ is a set of voxels representing a plane and the values of $L$ indicate different minerals or the background, e.g., we put $L(x) = 0$ if background was observed at $x \in H \cap W$, or $L(x) = 1$ if zinnwaldite was observed at $x \in H \cap W$. Furthermore, let $P_1, \ldots, P_n \subset W \subset \mathbb{Z}^3$ be the sets of voxels corresponding to the particles that are obtained by the particle-wise segmentation of the 3D XMT image and that hit the plane $H$, i.e., $P_k \cap H = \{x_1^{(k)}, \ldots, x_{\ell_k}^{(k)}\} \neq \emptyset$ for each $k = 1, \ldots, n$. Each particle $P_k$ gets the label

$$L(P_k) = \text{mode}(L(x_1^{(k)}), \ldots, L(x_{\ell_k}^{(k)})),$$

where the mode of the sample of labels $(L(x_1^{(k)}), \ldots, L(x_{\ell_k}^{(k)}))$ is the label that appears most often. Thus, we assign to each particle $P_k$ the mineral that is observed most frequently in the intersecting plane $H$. This means, for example, that a composite particle $P_k$ which consists of zinnwaldite, quartz and other minerals will be considered as a zinnwaldite particle if its main component in the intersection $P_k \cap H$ is zinnwaldite. Since we make this labeling only based on the mineralogical observation in the plane $H$ we assume stationarity of the mineralogical composition of the particles $P_k$, i.e., that the composition of a particle $P_k$ outside of the plane $H$ is adequately represented by $P_k \cap H$.

In our data the set of particles observed in the plane $H$ can be decomposed in three disjoint sets

$$\{P_1, \ldots, P_n\} = Z \cup Q \cup O$$

where $Z = \{P_k: L(P_k) = 1\}$ are the zinnwaldite particles and $Q = \{P_k: L(P_k) = 2\}$ are the quartz particles. The set $O = \{P_k: L(P_k) \geq 3\}$ contains the remaining particles that were observed in $H$. In the considered plane $H$ contains 861 particles, from which 342 belong to the set of zinnwaldite particles $Z$ and 462 to the set of quartz particles $Q$. Since the set $O$ contains only 57 particles, we disregard these observations from now on.

The goal of the present paper is to develop a method which allows us to predict the main mineralogical component $L(P)$ of particles when the additional planar SEM-EDS data is not available.

2.2 Particle characteristics

In Section 2.3 below, we will describe a method which allows us to determine the mineralogical characterization of particles solely based on XMT information. This will be done by a prediction model, which determines the mineralogical composition of a particle based on grayscale values and shape/size characteristics obtained from XMT data and its particle-wise segmentation. In order to adjust such a prediction model we use the ground truth information from a given planar SEM-EDS section, i.e., the sets of particles $Z$ and $Q$, for which we know that they consist in majority of zinnwaldite and quartz, respectively. After adjusting the prediction model we still used another spatially separated SEM-EDS section to validate the prediction model. In order to make such predictions, we first introduce some particle descriptors, by which a particle $P \subset \mathbb{Z}^3$ will be characterized.
2. MATERIALS AND METHODS

Modeling of Particle Properties

2.2.1 Size characteristics

Relevant size descriptors of a particle \( P' \subset \mathbb{R}^3 \), which we only observe on the grid as \( P \subset \mathbb{Z}^3 \), are the volume \( \nu_3(P') \) and the surface area \( a(P') \). It is clear that the volume can be estimated by counting voxels belonging to the set \( P \) and the surface area can be estimated by considering suitably defined voxel configurations, see Schladitz et al. (2006). We denote the estimated particle volume and surface area by \( \nu_3(P) \) and \( a(P) \) respectively.

For our prediction model we use the following estimator of the volume equivalent radius

\[
r(P) = \sqrt[3]{\frac{3}{4\pi}} \nu_3(P),
\]

which is in a one-to-one relationship with the volume. The surface area \( a(P) \) will not be directly regarded as a particle characteristic, but it is required for the sphericity factor which is a shape characteristic considered in the next section.

2.2.2 Shape characteristics

In the previous section we only considered the volume equivalent radius \( r \) as size characteristic, but this is not enough to reliably distinguish between zinnwaldite and quartz particles, since particles of similar sizes can have very different appearances. Therefore we examine some further characteristics for describing the shape of particles. One of the shape characteristics we consider is the sphericity factor

\[
s(P) = \frac{\sqrt[3]{36\pi \nu_2^2(P)}}{a(P)},
\]

which takes values in \([0, 1]\), where \( s(P') = 1 \) holds if \( P' \subset \mathbb{R}^3 \) is a sphere.

Another shape characteristic of interest is the convexity factor

\[
c(P) = \frac{\nu_3(P)}{\nu_3(q(P))},
\]

where \( q(P) \subset \mathbb{Z}^3 \) is the convex hull of \( P \) on the lattice \( \mathbb{Z}^3 \). Like the sphericity factor, the convexity factor takes values in \([0, 1]\), where \( c(P) = 1 \) holds if and only if \( P \) is convex on \( \mathbb{Z}^3 \). There is a causality between the sphericity factor \( s \) and the convexity factor \( c \) since more spherically shaped particles also have higher convexity factors. Yet, a convexity factor of 1 does not necessarily imply a large sphericity factor, since, e.g., ellipsoids have always a convexity factor of 1, yet their sphericity factor can be arbitrary close to 0.

Since we observe a large number of flat particles in our image data, we also consider characteristics that quantify elongation of particles. For each particle \( P \subset \mathbb{Z}^3 \) with the barycenter \( \vec{x} = (\vec{x}_1, \vec{x}_2, \vec{x}_3) \in \mathbb{R}^3 \), whose coordinates are given by

\[
\vec{x}_i = \frac{1}{\nu_3(P)} \sum_{x \in P} x_i \quad \text{for } i = 1, 2, 3,
\]

where \( x = (x_1, x_2, x_3) \), we consider the (positive-semidefinite) covariance matrix

\[
C = \left( \frac{1}{\nu_3(P)} \sum_{x \in P} (x_i - \overline{x}_i)(x_j - \overline{x}_j) \right)_{i,j=1,2,3}.
\]
The eigenvalues $0 \leq \lambda_1 \leq \lambda_2 \leq \lambda_3$ of $C$, where we assume that $\lambda_3 > 0$, are closely related to the axis lengths of the best fitting ellipsoid $\varepsilon(P)$ corresponding to the particle $P$, i.e., the axis lengths $a_1, a_2, a_3$ of $\varepsilon(P)$ are given by $a_i = \gamma \sqrt{\lambda_i}$ for $i = 1, 2, 3$, where $\gamma$ is some scaling factor. The elongation factor $e(P)$ of the particle $P$ is then given by

$$e(P) = \frac{a_1}{a_3} = \sqrt{\frac{\lambda_1}{\lambda_3}}.$$  \hspace{1cm} (8)

Note that, like the previously considered characteristics, the elongation factor $e(P)$ is normalized and takes values in $[0, 1]$, where values close to 0 indicate elongated particles like rod- or plate-shaped particles. By additionally analyzing the relationship between $\lambda_2$ and $\lambda_3$ it is possible to distinguish between rods and plates, but since we observed no rod-like particles in our data set, this was not necessary in the present paper.

### 2.2.3 Grayscale characteristics

The XMT data does not only provide information about the 3D morphology of particles. In Furat et al. (2018) we have seen that the grayscale values of XMT images contain substantial information about the mineralogical composition of particles. Recall that the grayscale values of a particle $P = \{x_1, \ldots, x_n\} \subset W$ are given by $y_1 = I(x_1), \ldots, y_n = I(x_n) \in [0, 1]$, where $I$ is the XMT image. Furthermore, we assume without loss of generality that the grayscale values are ordered, i.e., $y_1 \leq y_2 \leq \ldots \leq y_n$. This allows us to define the following grayscale characteristics.

A natural choice could be the mean

$$\bar{y} = \frac{1}{n} \sum_{k=1}^{n} y_k.$$  \hspace{1cm} (9)

But, since particles often have some imperfections, like small regions of high density, see Figure 2, the mean is not well suited for representing the dominant grayscale value. Therefore, we used the more robust median

$$m(P) = y_{0.5}$$  \hspace{1cm} (10)

as grayscale characteristic, where the median is defined by the empirical quantiles of the sample

$$y_p = \begin{cases} y_{[np]+1}, & \text{if } np \in \mathbb{N}, \\ (y_{[np]} + y_{[np]+1})/2, & \text{if } np \notin \mathbb{N}, \end{cases}$$  \hspace{1cm} (11)

for $p = 0.5$. A natural choice for measuring the variability of grayscale values of a particle could be the sample standard deviation

$$\sigma = \sqrt{\frac{1}{n-1} \sum_{k=1}^{n} (y_k - \bar{y})^2}.$$  \hspace{1cm} (12)

Similarly to the mean, the standard deviation is not a robust particle descriptor. Therefore we use the interquartile range

$$iqr(P) = y_{0.75} - y_{0.25}$$  \hspace{1cm} (13)

to characterize the variability of the grayscale values of the dominant grayscale region of the particle. Further characteristics that describe the texture of particles (Shapiro and Stockman, 2001) will be investigated in a forthcoming paper.
2.3 Stochastic modeling of particle characteristics

In this section we fit parametric distributions to the particle characteristics discussed above. Moreover, we fit parametric distributions separately for zinnwaldite and quartz particles, since the mineralogical composition of particles that intersect with the SEM-EDS data is known by the labeling considered in (1). In Figure 3 the fitted one-dimensional distributions of the median grayscale value for both zinnwaldite and quartz are visualized, where we can see that both minerals have quite distinct distributions. Still, there are overlapping regions which make a classification merely based on these distributions difficult. To be more precise, let \( f_{zm} \), \( f_{qm} \) be the fitted probability density functions of the particle-wise median grayscale values of zinnwaldite and quartz, respectively. Using a likelihood approach for classification, we say that a particle \( P \subset \mathbb{Z}^3 \) with median grayscale value \( x = m(P) \in [0,1] \) is mainly composed of zinnwaldite if

\[
f_{zm}(x) > f_{qm}(x),
\]

otherwise we say that \( P \) is a quartz particle. The probability of misclassifying zinnwaldite particles is then given by

\[
P\left(f_{zm}(X) \leq f_{qm}(X)\right) = \int_0^1 f_{zm}(x) \mathbf{1}_{f_{zm}(x) \leq f_{qm}(x)} \, dx,
\]

where \( X \) is a random median grayscale value with probability density \( f_{zm} \) and \( \mathbf{1}_{f_{zm}(x) \leq f_{qm}(x)} \) denotes the indicator function, i.e.,

\[
\mathbf{1}_{f_{zm}(x) \leq f_{qm}(x)} = \begin{cases} 
1, & \text{if } f_{zm}(x) \leq f_{qm}(x), \\
0, & \text{if } f_{zm}(x) > f_{qm}(x). 
\end{cases}
\]

This misclassification probability corresponds to the blue area in Figure 3 and has the size of 0.19. Furthermore, Table 1 shows a confusion matrix of the prediction rule given in (14), where we can see that 73 out of 342 zinnwaldite particles in the SEM-EDS plane were wrongly classified as quartz particles.
2. MATERIALS AND METHODS

Modeling of Particle Properties

Figure 3: One-dimensional probability densities of the median grayscale value of zinnwaldite (blue curve) and quartz (red curve) particles. The empirical densities were fitted with Beta distributions, see Section 2.3.1. The orange area provides the probability of classifying a quartz particle as zinnwaldite, whereas the blue area gives the probability of classifying zinnwaldite as quartz.

Table 1: Confusion matrix for predicting the particle composition based on one-dimensional densities of the median grayscale value, where the particles observed in the SEM-EDS data have been used for adjusting the prediction model.

<table>
<thead>
<tr>
<th>predicted zinnwaldite</th>
<th>zinnwaldite</th>
<th>quartz</th>
</tr>
</thead>
<tbody>
<tr>
<td>predicted quartz</td>
<td>269</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>73</td>
<td>436</td>
</tr>
</tbody>
</table>

In order to reduce the misclassification probabilities we additionally consider multidimensional distributions of vectors of particle characteristics. For example, in Figure 4 the joint probability density of the sphericity factor and the median grayscale value is visualized for zinnwaldite (left) and quartz (right) particles. We can see that these two-dimensional distributions are much more distinct, in comparison to the one-dimensional distributions of Figure 3. In accordance with this, the consideration of bivariate probability densities leads to a remarkable drop of the zinnwaldite misclassification probability from 0.19 to 0.06, where we used a two-dimensional analogue of the decision rule given in (14), see also (38) for the general definition of a multidimensional decision rule. Note that the number of wrongly classified zinnwaldite particles dropped from 73 to 27, see Table 2, whereas the number of wrongly classified quartz particles only increased slightly. Thus, by considering multidimensional probability distributions of particle characteristics we can achieve better prediction results. In the following we describe in detail how we constructed such two- or even higher-dimensional probability distributions and their corresponding decision rules.
2. MATERIALS AND METHODS

Modeling of Particle Properties

Figure 4: Joint probability density of sphericity factor and median grayscale value for zinnwaldite (left) and quartz (right) particles.

Table 2: Confusion matrix for predicting the particle composition based on joint densities of the sphericity factor and the median grayscale value, where the particles observed in the SEM-EDS data have been used for adjusting the prediction model.

<table>
<thead>
<tr>
<th></th>
<th>zinnwaldite</th>
<th>quartz</th>
</tr>
</thead>
<tbody>
<tr>
<td>predicted zinnwaldite</td>
<td>315</td>
<td>33</td>
</tr>
<tr>
<td>predicted quartz</td>
<td>27</td>
<td>429</td>
</tr>
</tbody>
</table>

2.3.1 Modeling of single particle characteristics

We now fit parametric probability distributions to the one-dimensional particle characteristics described in Section 2.2. To be precise, we first fit parametric distributions to the volume equivalent radius, sphericity factor, convexity factor, elongation factor, median grayscale value and the interquartile range for the 3D zinnwaldite particles that intersect with the SEM-EDS plane. The chosen families of parametric distributions and the corresponding parameters that were determined to model the distributions of these characteristics are listed in Table 3. Note that we used gamma- and beta-distributions, whose probability density functions are given by

\[
f_{\text{gamma}}(x) = \frac{1}{\theta^k \Gamma(k)} x^{k-1} \exp\left(-\frac{x}{\theta}\right) \mathbb{1}_{x>0} \tag{17}
\]

and

\[
f_{\text{beta}}(x) = \frac{\Gamma(\alpha) \Gamma(\beta)}{\Gamma(\alpha + \beta)} x^{\alpha-1} (1-x)^{\beta-1} \mathbb{1}_{x \in [0,1]}, \tag{18}
\]

respectively, where \(k, \theta, \alpha, \beta > 0\) are some model parameters and the gamma function \(\Gamma: [0, \infty) \rightarrow [0, \infty)\) is defined by the integral

\[
\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx \quad \text{for each } \alpha > 0. \tag{19}
\]
2. MATERIALS AND METHODS

Modeling of Particle Properties

Figure 5: Histograms (blue) and fitted probability densities of zinnwaldite particle characteristics for a) volume equivalent radius, b) sphericity factor, c) convexity factor, d) elongation factor, e) median grayscale value, f) interquartile range, see also Table 3.

For modeling the volume equivalent radius we used the gamma distribution. Note that there are also other families of distributions which result in a good fit, like the log-normal distribution. The remaining characteristics were modeled with beta distributions, since these particle characteristics have only values in the interval \([0,1]\) which coincides with the support of the beta distribution. A visual comparison with the histograms of particle characteristics is given in Figure 5. In order to determine the model parameters \((k,\theta)\) for the gamma- and \((\alpha,\beta)\) for the beta distribution) for these one-dimensional fits we used maximum likelihood estimators (Held and Bové, 2014). Analogously, the one-dimensional distributions of single particle characteristics were fitted for quartz particles, see Figure 6 and Table 3.

Table 3: Parameters of fitted distributions.

<table>
<thead>
<tr>
<th>characteristic</th>
<th>distribution</th>
<th>zinnwaldite</th>
<th>quartz</th>
</tr>
</thead>
<tbody>
<tr>
<td>volume equivalent radius</td>
<td>gamma</td>
<td>(k = 7.16, \theta = 6.84)</td>
<td>(k = 6.84, \theta = 9.8)</td>
</tr>
<tr>
<td>sphericity factor</td>
<td>beta</td>
<td>(\alpha = 7.17, \beta = 7.05)</td>
<td>(\alpha = 10.73, \beta = 4.63)</td>
</tr>
<tr>
<td>convexity factor</td>
<td>beta</td>
<td>(\alpha = 12.61, \beta = 2.97)</td>
<td>(\alpha = 17.35, \beta = 2.73)</td>
</tr>
<tr>
<td>elongation factor</td>
<td>beta</td>
<td>(\alpha = 2.00, \beta = 4.73)</td>
<td>(\alpha = 2.90, \beta = 2.74)</td>
</tr>
<tr>
<td>median grayscale value</td>
<td>beta</td>
<td>(\alpha = 14.12, \beta = 18.32)</td>
<td>(\alpha = 74.95, \beta = 166.85)</td>
</tr>
<tr>
<td>interquartile range</td>
<td>beta</td>
<td>(\alpha = 3.78, \beta = 38.54)</td>
<td>(\alpha = 4.52, \beta = 112.11)</td>
</tr>
</tbody>
</table>

2.3.2 Modeling of multidimensional particle characteristics

When modeling the joint distribution of independent random particle characteristics \(X, Y\) with probability density functions \(f_X\) and \(f_Y\), respectively, the joint distribution is immediately given by \(f_{(X,Y)}(x,y) = f_X(x)f_Y(y)\). However, this approach is not available
for correlated characteristics, as it is the case for our data, where, for example, we obtained a correlation coefficient of 0.36 for the sphericity factor and median grayscale value of zinnwaldite particles, see Table 4. A rather easy way for describing correlated characteristics are multivariate normal distributions, see Anderson (2003). However, the marginals of such distributions are always normal distributions, which is not the case for our data, see Figures 5 and 6, where we used gamma- and beta-distributions to model the (one-dimensional) distributions of single particle characteristics. For instance, the histograms of the characteristics in Figures 5 c) and d) are clearly skewed, whereas the density functions of normal distributions are symmetric around their mean values.

Table 4: Correlation coefficients of particle characteristics for zinnwaldite.

<table>
<thead>
<tr>
<th></th>
<th>r</th>
<th>s</th>
<th>c</th>
<th>e</th>
<th>m</th>
<th>iqr</th>
</tr>
</thead>
<tbody>
<tr>
<td>r</td>
<td>1.0000</td>
<td>0.1450</td>
<td>0.0695</td>
<td>0.3443</td>
<td>0.4939</td>
<td>0.2825</td>
</tr>
<tr>
<td>s</td>
<td>0.1450</td>
<td>1.0000</td>
<td>0.5936</td>
<td>0.4400</td>
<td>0.3636</td>
<td>0.2632</td>
</tr>
<tr>
<td>c</td>
<td>0.0695</td>
<td>0.5936</td>
<td>1.0000</td>
<td>-0.0663</td>
<td>0.3606</td>
<td>0.1267</td>
</tr>
<tr>
<td>e</td>
<td>0.3443</td>
<td>0.4400</td>
<td>-0.0663</td>
<td>1.0000</td>
<td>0.2155</td>
<td>0.2771</td>
</tr>
<tr>
<td>m</td>
<td>0.4939</td>
<td>0.3636</td>
<td>0.3606</td>
<td>0.2155</td>
<td>1.0000</td>
<td>0.2276</td>
</tr>
<tr>
<td>iqr</td>
<td>0.2825</td>
<td>0.2632</td>
<td>0.1267</td>
<td>0.2771</td>
<td>0.2276</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

Thus, in order to describe the joint distribution of vectors of particle characteristics, whose marginal distributions are given by the one-dimensional distributions we fitted in Section 2.3.1, see Table 3, we consider a more general approach, using so-called Archimedean copulas, see Nelsen (2006).
2.3.3 Basic ideas of the copula approach

To make the paper more self-contained, we first explain some basic ideas of the copula approach, which will be used to construct multivariate probability distributions with non-normal marginals.

Therefore, let \( f_1, \ldots, f_d : \mathbb{R} \to [0, \infty) \) denote the one-dimensional probability densities of particle characteristics for which we want to determine the \( d \)-dimensional joint probability density \( f : \mathbb{R}^d \to [0, \infty) \) whose marginal densities are given by \( f_1, \ldots, f_d \). In order to construct the multivariate density \( f \) we need to consider, for technical reasons, the one-dimensional (cumulative) distribution functions \( F_1, \ldots, F_d : \mathbb{R} \to [0, 1] \) which are given by

\[
F_i(x) = \int_{-\infty}^{x} f_i(y) \, dy \quad \text{for } i = 1, \ldots, d. \tag{20}
\]

Note that \( F_i(x) \) denotes the probability that the particle characteristic described by the density \( f_i \) does not exceed the value \( x \in \mathbb{R} \) and that the density can be obtained from the distribution function \( F_i \) by

\[
f_i(x) = \frac{d}{dx} F_i(x) \quad \text{for } i = 1, \ldots, d. \tag{21}
\]

Analogously, the \( d \)-dimensional density \( f \) has a cumulative distribution function \( F : \mathbb{R}^d \to [0, 1] \) which is given by

\[
F(x) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_d} f(y_1, \ldots, y_d) \, dy_d \cdots dy_1 \quad \text{for } x = (x_1, \ldots, x_d) \in \mathbb{R}^d. \tag{22}
\]

The joint density \( f \) is then obtained by the \( d \)-fold partial derivative

\[
f(x) = \frac{\partial^d}{\partial x_1 \cdots \partial x_d} F(x) \quad \text{for } x = (x_1, \ldots, x_d) \in \mathbb{R}^d. \tag{23}
\]

In order to show the connection between the multivariate distribution function \( F \) and the notion of copulas we begin with the definition of the latter. A \( d \)-dimensional copula is a multivariate cumulative distribution function \( K : \mathbb{R}^d \to [0, 1] \), whose one-dimensional marginal distributions are uniform distributions on the interval \([0, 1]\). For example the marginal cumulative distribution function \( K_1 : \mathbb{R} \to [0, 1] \) of the first component is given by

\[
K_1(x_1) = \lim_{x_2, \ldots, x_d \to \infty} K(x_1, x_2, \ldots, x_d) = \begin{cases} 0, & \text{if } x_1 < 0, \\ x_1, & \text{if } x_1 \in [0, 1], \\ 1, & \text{if } x_1 > 1. \end{cases} \tag{24}
\]

Copulas are of special interest because of Sklar’s theorem (Nelsen, 2006). It states that \( F \) is a \( d \)-dimensional cumulative distribution function with marginal distribution functions \( F_1, \ldots, F_d \) if and only if there is a copula function \( K \) such that

\[
F(x) = K(F_1(x_1), \ldots, F_d(x_d)) \quad \text{for } x = (x_1, \ldots, x_d) \in \mathbb{R}^d. \tag{25}
\]

This means that every multivariate distribution function \( F \) can be represented by its marginals \( F_1, \ldots, F_d \) and a copula function \( K \). Thus, when modeling the joint distribution function \( F \) of random vectors whose marginal distributions are given by \( F_1, \ldots, F_d \) it suffices to model the copula \( K \). Note that there are numerous classes of copula functions, for which many of them do not have an analytical representation. In the present paper we limit the search for an appropriate copula function to so-called Archimedean copulas.
2.3.4 Archimedean copulas and differential variant of Sklar’s theorem

For that purpose, let \( \varphi : [0, 1] \to [0, \infty] \) be a continuous, strictly decreasing and convex function with \( \varphi(1) = 0 \) and \( \varphi(0) = \infty \). Note that \( \varphi \) is called an Archimedean generator.

Then, it can be shown that the function \( K : [0, 1]^d \to [0, 1] \) given by

\[
K(u) = \varphi^{-1}(\varphi(u_1) + \cdots + \varphi(u_d)) \quad \text{for } u = (u_1, \ldots, u_d) \in [0, 1]^d
\]

(26)
can be uniquely extended to a copula on \( \mathbb{R}^d \), i.e., a function possessing the properties mentioned in Section 2.3.3. It is called an Archimedean copula (Nelsen, 2006). With the copula function \( K \) given in (26) which still depends on some abstract function \( \varphi \) we can construct a \( d \)-dimensional distribution function \( F \) by means of (25) which has the marginal distribution functions \( F_1, \ldots, F_d \). Since, finally, we are interested in the joint density \( f \) of particle characteristics, applying (23) to (25) we get that

\[
f(x) = f_1(x_1) \cdots f_d(x_d) k(F_1(x_1), \ldots, F_d(x_d)) \quad \text{for } x = (x_1, \ldots, x_d) \in \mathbb{R}^d,
\]

(27)
which can be seen as a differential variant of Sklar’s theorem given in (25), where the function \( k : \mathbb{R}^d \to [0, \infty) \) is the \( d \)-fold derivative

\[
k(u) = \frac{\partial^d}{\partial u_1 \cdots \partial u_d} K(u) \quad \text{for } u = (u_1, \ldots, u_d) \in \mathbb{R}^d.
\]

(28)

Recall that the \( d \)-dimensional probability density \( f \) given by (27) has the marginal densities \( f_1, \ldots, f_d \) and, besides this, depends on the choice of a suitable copula function \( K \) or, equivalently, the choice of its derivative \( k \). Thus, the task in modeling multidimensional probability distributions for given (one-dimensional) marginal densities is the determination of a suitable copula function \( K \) which describes the correlation structure of the one-dimensional marginals. Since in the present paper we only consider Archimedean copulas, which are defined by their generator \( \varphi \), this task is equivalent to finding a suitable Archimedean generator.

Note that, for example, the generator given by \( \varphi(u) = -\log(u) \) with the inverse \( \varphi^{-1}(u) = e^{-u} \) induces the copula function

\[
K(u) = u_1 \cdots u_d \quad \text{for } u = (u_1, \ldots, u_d) \in [0, 1]^d.
\]

(29)

For this example the resulting joint density \( f \) in (27) is given by

\[
f(x) = f_1(x_1) \cdots f_d(x_d) \quad \text{for } x = (x_1, \ldots, x_d) \in \mathbb{R}^d,
\]

(30)
which means that the random particle characteristics described by the marginal densities \( f_1, \ldots, f_d \) would be independent. However, as mentioned above, this is not the case for the characteristics considered in the present paper. Thus, the so-called independence copula given in (29) is not suitable for modeling the joint density \( f \).

2.3.5 Parametric family of Archimedean generators

Instead of analyzing single generator functions \( \varphi \) one by one, we can instead consider parametric families \( \{ \varphi_\theta : \theta \in \Theta \} \) of generators. For such families the function \( \varphi_\theta \) is an Archimedean generator for each parameter \( \theta \in \Theta \), where \( \Theta \) is some set of admissible parameters. This has the advantage that we can consider a whole range of generators for
2. MATERIALS AND METHODS

Modeling of Particle Properties

which it is rather easy to determine an optimal choice of a generator among the family
\{\varphi_{\theta} : \theta \in \Theta\} for modeling the joint density \( f \).

A parametric family of generators that provided a good fit for our data was the Ali-
Mikhail-Haq generator, see Nelsen (2006), which is given by

\[
\varphi_{\theta}(u) = \log \left( \frac{1 - \theta (1 - u)}{u} \right),
\]

for each \( u \in [0, 1] \) and some \( \theta \in \Theta = (-1, 1) \). Fitting a model from this parametric
family of generators is equivalent to determining an optimal parameter \( \hat{\theta} \). This problem

The family of generators \{\varphi_{\theta} : \theta \in \Theta\} induces a family of copulas \{K_{\theta} : \theta \in \Theta\} given
by

\[
K_{\theta}(u) = \varphi_{\theta}^{-1}(\varphi_{\theta}(u_1) + \cdots + \varphi_{\theta}(u_d)) \quad \text{for } u = (u_1, \ldots, u_d) \in [0, 1]^d,
\]

for which the optimal parameter \( \hat{\theta} \) has to be determined. Since \( K_{\theta} \) is differentiable we

and the corresponding \( d \)-dimensional distribution function \( F_{\theta} \) which is given by (25) has the probability density function

\[
f_{\theta}(x) = f_1(x_1) \cdots f_d(x_d) k_{\theta}(F_1(x_1), \ldots, F_d(x_d)) \quad \text{for } x = (x_1, \ldots, x_d) \in \mathbb{R}^d,
\]

where \( f_1, \ldots, f_d \) and \( F_1, \ldots, F_d \) are the one-dimensional densities and distribution functions fitted in Section 2.3.1 to the image data considered in the present paper.

Note that, for a sample of \( \ell \in \mathbb{N} \) observations \( x^{(1)}, \ldots, x^{(\ell)} \in \mathbb{R}^d \) of the considered \( d \)-dimensional vector of particle characteristics, the so-called log-likelihood function is given by

\[
\log \mathcal{L}(\theta|x^{(1)}, \ldots, x^{(\ell)}) = \sum_{k=1}^{\ell} \log \left( f_{\theta}(x^{(k)}) \right) \quad \text{for } \theta \in (-1, 1).
\]

This leads to the maximum-likelihood estimator

\[
\hat{\theta} = \arg \max_{\theta \in (-1, 1)} \log \mathcal{L}(\theta|x^{(1)}, \ldots, x^{(\ell)}),
\]

which can be considered as the optimal choice for the parameter \( \theta \) when the sample
\( x^{(1)}, \ldots, x^{(\ell)} \) was observed. The fitted joint density \( f \) is then given by \( f = f_{\hat{\theta}} \).

2.3.6 Application to particle characteristics

In the context considered in the present paper, the estimation procedure described above can be applied in the following way. Recall that in Section 2.3.1 we determined, for each zinnwaldite particle \( P_1, \ldots, P_\ell \in Z \) that hits the SEM-EDS plane, the 6-dimensional vector of particle characteristics \( x^{(j)} = (r(P_j), s(P_j), c(P_j), e(P_j), m(P_j), iqr(P_j)) \), where the components of \( x^{(j)} \) are the volume equivalent radius, the sphericity factor, the convexity factor, the elongation factor, the median grayscale value and the interquartile grayscale range of particle \( P_j \). For each of these six characteristics the parameters of the
corresponding one-dimensional probability densities $f^z_r, f^z_c, f^z_s, f^z_m, f^z_{iqr}$ for zinnwaldite particles are given in Table 3. Using (32)-(35) we can compute the log-likelihood function $\log L$, see Figure 7, and obtain the copula parameter $\hat{\theta} = 0.69$. By means of (34) we then get the joint probability density function of zinnwaldite particle characteristics

$$f^z(x_1, \ldots, x_6) = f^z_r(x_1) \ldots f^z_{iqr}(x_6) k^z_{\hat{\theta}}(F^z_r(x_1), \ldots, F^z_{iqr}(x_6)),$$

where $F^z_r$ is the distribution function corresponding to the density $f^z_r$, and so on. Analogously, we determined the copula parameter $\hat{\theta} = 0.47$ and thus the multidimensional density function $f^q$ for the characteristics of quartz particles.

3 Results and Discussion

3.1 Copula-based classification model

3.1.1 Decision rule

At the beginning of Section 2.3 we showed a way how the mineralogical composition of particles could be predicted on the basis of one-dimensional probability distributions of single particle characteristics introduced in Section 2.2. However, in Figure 3 and Table 1 we have seen that one-dimensional distributions do not suffice for making reliable decisions, whereas Figure 4 and Table 2 indicate that the predictive power of our decision rule can increase when we consider bivariate vectors of characteristics. Moreover, in Section 2.3.6 we showed how the multidimensional density functions $f^z$ and $f^q$ for 6 characteristics of zinnwaldite and quartz particles can be determined. Then, analogously to the one-dimensional case, it is possible to predict the composition of a particle $P \subset \mathbb{Z}^3$ observed in the XMT image data, using the following decision rule: If

$$f^z(r(P), s(P), c(P), e(P), m(P), iqr(P)) > f^q(r(P), s(P), c(P), e(P), m(P), iqr(P)),$$

then we assume that $P$ is mainly composed of zinnwaldite, otherwise of quartz. In other words, Equation (38) states that a particle $P$ with the vector of characteristics $x = (r(P), s(P), c(P), e(P), m(P), iqr(P))$ is classified as a zinnwaldite particle, if the configuration of characteristics $x$ is more likely to occur according to the distribution of characteristics of zinnwaldite particles given by $f^z$. 

Figure 7: Log-likelihood functions for given marginal density functions, in order to estimate the copula parameter $\theta$ for zinnwaldite (left) and quartz (right).
3. RESULTS AND DISCUSSION

Modeling of Particle Properties

Table 5: Confusion matrix of the copula-based decision rule given in (38), where the particles observed in the SEM-EDS data have been used for adjusting the prediction model.

<table>
<thead>
<tr>
<th>predicted zinnwaldite</th>
<th>zinnwaldite</th>
<th>quartz</th>
</tr>
</thead>
<tbody>
<tr>
<td>predicted zinnwaldite</td>
<td>333</td>
<td>25</td>
</tr>
<tr>
<td>predicted quartz</td>
<td>9</td>
<td>437</td>
</tr>
</tbody>
</table>

Table 5 shows that the decision rule given in (38) is rather accurate, i.e., only 2.7% of zinnwaldite particles are wrongly classified as quartz particles and 5.7% of the quartz particles are wrongly classified as zinnwaldite particles. In comparison to the decision rule which only used the sphericity factor and the median grayscale value, see Table 2, the number of wrongly classified zinnwaldite particles dropped from 27 to 9. This is a significant improvement compared to the prediction based on two characteristics shown in Table 2. Thus, we assume that the prediction results can become even better when we consider more than six characteristics. However, both the computational as well as the model complexity increase when we consider additional characteristics.

3.1.2 Kernel density estimation

Note that the decision rule proposed in (38) solely requires the multivariate probability densities \( f_z \) and \( f_q \) of vectors of characteristics for zinnwaldite and quartz particles, respectively. In Section 2.3.2 these were determined using a parametric family of copulas. Alternatively, such densities could be estimated in a nonparametric way, using so-called kernel density estimators (Scott, 2015). To be precise, from a sample \( x^{(1)}, \ldots, x^{(\ell)} \in \mathbb{R} \) of a one-dimensional characteristic of zinnwaldite particles a non-parametric density \( \hat{f}_z \), can be estimated by

\[
\hat{f}_z(x) = \frac{1}{\ell} \sum_{j=1}^{\ell} \kappa\left(\frac{x - x^{(j)}}{h}\right) \quad \text{for } x \in \mathbb{R},
\]

(39)

where \( h > 0 \) is called the bandwidth and \( \kappa: \mathbb{R} \rightarrow [0, \infty) \) is a kernel function, e.g.,

\[
\kappa(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \quad \text{for } x \in \mathbb{R}.
\]

(40)

Analogously, it would be possible to estimate non-parametric multivariate densities for vectors of particle characteristics.

At first glance such an approach would make the parametric fitting of one-dimensional distributions in Section 2.3.1 and the parametric copula calculus of Section 2.3.2 obsolete. However, due to the “curse of dimensionality”, as mentioned by Scott (2015), the required sample size for accurate kernel density estimation increases exponentially with the increasing dimension of vector data. For example, for a two-dimensional kernel density estimation one would need a sample size of 258 in order to reach the same accuracy as an one-dimensional density estimated from a sample of size 50. Moreover, for a three-dimensional kernel density estimation with the same accuracy one would need already a sample size of 1126. Since we only observed 342 zinnwaldite and 462 quartz particles in the considered SEM-EDS plane, the approach of using kernel density estimators for multivariate density functions \( f_z \) and \( f_q \) is impractical. Furthermore, kernel density estimators lead to rather complex density functions, whereas, the approach of using parametric copulas reduces the complexity of a distribution to a few parameters, e.g., the six-dimensional densities \( f_z \) and \( f_q \) considered in (38) are each described with only 13 parameters.
3.1.3 Model validation

To validate the predictive capability of our model, we used an additional planar SEM-EDS data set measured at a spatially different location of the 3D sample, see red plane in Figure 1a. For particles that hit this second plane we determined the particle characteristics and predicted the composition of the particles using the method described in Section 3.1.1. Due to the SEM-EDS data the true composition of these particles is known. A comparison between the ground truth and the prediction can be seen in Table 6, where the predictions were again rather good. Since in the latter case the mineralogical information of particles that hit the second SEM-EDS plane was not used for calibrating the prediction model, we can assume, based on the results of Table 6, that the predictions will remain accurate for each particle of the 3D XMT image.

Table 6: Confusion matrix of the copula-based decision rule given in (38) for particles observed in spatially separated SEM-EDS data that was not used for calibrating the prediction model.

<table>
<thead>
<tr>
<th>predicted</th>
<th>zinnwaldite</th>
<th>quartz</th>
</tr>
</thead>
<tbody>
<tr>
<td>predicted zinnwaldite</td>
<td>204</td>
<td>19</td>
</tr>
<tr>
<td>predicted quartz</td>
<td>12</td>
<td>244</td>
</tr>
</tbody>
</table>

4 Conclusions & Outlook

We presented a method for the mineralogical characterization of particles observed in 3D XMT image data which mainly consists of the minerals zinnwaldite and quartz. The grayscale value $I(x)$ of a voxel $x \in W \subset Z^3$ in the 3D XMT image already provides some information about the mass density of the observed mineral at the location $x$. However, this information does not suffice for reliably distinguishing between zinnwaldite and quartz particles. Therefore, we additionally considered the morphology of particles to characterize them. The proposed prediction model can then characterize particles based on their 3D morphology and grayscale values extracted solely from XMT data.

For the calibration of the prediction model we had to localize a 2D SEM-EDS data set in the 3D XMT image. The SEM-EDS data provided a mineralogical characterization of particles in a planar cross-section of the sample, such that we know the 3D morphology and the composition of particles that intersect with this cross-section. In order to extract single 3D particles that hit the planar cross-section, we computed a particle-wise segmentation of the 3D image data. This allowed us to determine for each particle its vector of characteristics which is relevant to distinguish between zinnwaldite and quartz particles. Among the considered characteristics are the median grayscale value of particles, but also particle size and shape characteristics like the volume equivalent radius and the sphericity factor.

This resulted in vectors of characteristics for each particle. Since the mineralogical composition of particles that hit the SEM-EDS plane is known, we assigned the corresponding vectors of characteristics to the mineral the particle is mostly composed of. Therefore, we had a set of vectors of characteristics which corresponded to zinnwaldite particles and analogously another set of such vectors corresponding to quartz particles. The proposed prediction model requires the joint densities $f_z$ and $f_q$ of these characteristics for particles composed of zinnwaldite and quartz, respectively. The determination of these multidimensional densities with kernel density estimation would be difficult,
since such an approach requires rather large data sets. Therefore, we fitted parametric copulas to each of the two data sets. This entailed fitting one-dimensional parametric families of distributions to each considered characteristic (separately for zinnwaldite and quartz particles). In a second step these one-dimensional densities were combined to obtain multidimensional densities. Since the considered characteristics were correlated the multidimensional densities had to reflect the correlation structure. We achieved this with the help of parametric Archimedean copulas.

The resulting decision rule characterizes a particle with a vector of characteristics \( x \in \mathbb{R}^d \) as zinnwaldite particle if \( f_z(x) > f_q(x) \). We observed that this decision rule became more accurate when higher dimensional vectors of characteristics were considered. However, there was already a significant drop of the misclassification percentage when we considered two-dimensional characteristics instead of just considering single (i.e. one-dimensional) particle characteristics. This error was reduced even further when we considered vectors of six characteristics. To validate the prediction model we used additional SEM-EDS data at a spatially different location than the SEM-EDS data set that was used to adjust the model. This allowed us to compare the predictions of the model with the ground truth obtained by SEM-EDS. Since we observed that the prediction model was quite accurate in this validation step, we can assume that it can reliably distinguish between zinnwaldite and quartz particles in the entire XMT image – even for areas where no SEM-EDS data is available.

In a forthcoming study we will use the copula-based modeling considered in the present paper to quantify the success of particle separation processes. To be precise, assume that the multidimensional density \( f \) of characteristics of a mixture of zinnwaldite and quartz particles is given by the convex-combination

\[
f = (1 - p)f_z + pf_q, \tag{41}
\]

for some \( p \in [0, 1] \). When the \( p \) value drops to 0 after a separation process we can say that the zinnwaldite particles are separated from quartz particles. However, it is possible that the distribution of zinnwaldite particle characteristics itself changes significantly during separation. For example, if the separation process only extracts small zinnwaldite particles, the multidimensional density of zinnwaldite characteristics after separation would differ from the original density \( f_z \), prior to separation. In order to track and compare these changes in the densities, it is useful to represent the multidimensional densities \( f_z \) prior and after separation with parametric copulas since they are described by only a few parameters. This complexity reduction is another advantage of parametric copulas, since it allows us to compare complicated multidimensional distributions by simply comparing the parameters of the distributions.

5 Acknowledgements

The financial support of the German Research Foundation (DFG) for funding the X-ray microscope (INST267/129-1) as well as the research projects (PE1160/22-1 and SCHM997/27-1) within the priority program SPP 2045 “Highly specific and multidimensional fractionation of fine particle systems with technical relevance” is gratefully acknowledged. The authors would like to thank Sabine Gilbricht for her work at the MLA and Roland Würkert for preparing the epoxy blocks.
References


