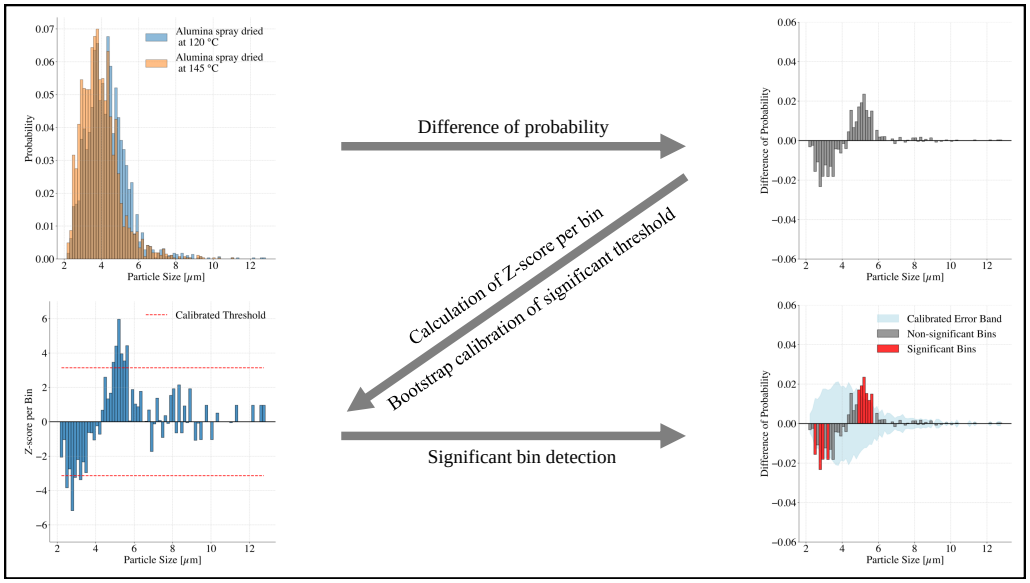
Graphical Abstract

A bin-resolved statistical framework for comparing particle size distributions

Rahul Mitra, Lukas Fuchs, Ralf Ditscherlein, Volker Schmidt, Urs Alexander Peuker



Highlights

A bin-resolved statistical framework for comparing particle size distributions

Rahul Mitra, Lukas Fuchs, Ralf Ditscherlein, Volker Schmidt, Urs Alexander Peuker

- Introduces a bin-resolved framework for comparing particle size distributions.
- Localizes statistically significant size differences between particle populations.
- Uses parametric bootstrap calibration for approximate simultaneous error control.
- Combines global hypothesis testing with local bin-wise statistical inference.

A bin-resolved statistical framework for comparing particle size distributions

Rahul Mitra[a,*], Lukas Fuchs[b], Ralf Ditscherlein[a], Volker Schmidt[b], Urs Alexander Peuker[a]

[a] *Institute of Mechanical Process Engineering and Mineral Processing, Technische Universität Bergakademie Freiberg, Agricolastr. 1, Freiberg, 09599, Saxony, Germany*
[b] *Institute for Stochastics, Universität Ulm, Helmholtzstr. 18, Ulm, 89069, Baden-Württemberg, Germany*

Abstract

Reliable comparison of particle size distributions (PSDs) is crucial for monitoring, optimizing, and designing powder-based processes in industries such as spray drying, granulation, milling, classification, and physical separation. Subtle variations in PSDs often indicate changes in process stability, product quality, or material functionality. However, conventional methods such as visual histogram comparison or global statistical tests, such as the Pearson χ^2 test, lack calibration or do not pinpoint the size classes (also referred to as "bins") in which the differences occur. To overcome these limitations, we introduce the Difference Significance Profile (DSP), a statistically grounded and practically interpretable framework for comparing two PSDs at the bin level. The framework quantifies the observed differences between the corresponding bins and identifies those size classes in which the differences are too large to be plausibly explained by random sampling after simultaneous calibration across all bins. It uses a pooled-variance estimator and a bootstrap-calibrated simultaneous threshold that approximately controls the family-wise error rate (FWER) at a predetermined significance level, thereby reducing false-positive detections. We demonstrate the framework on number-weighted PSDs measured using in-line imaging of spray-dried alumina, quantify the temperature-induced PSD changes, and illustrate the behavior of DSP in a negative-control comparison under nominally identical operating conditions. In general, DSP provides a quantitative and practically interpretable tool that supports localized comparison of particle size distributions.

Keywords: Particle size distribution, Bin-resolved analysis, Statistical inference, Bootstrap calibration, Simultaneous hypothesis testing

1. Introduction

Comparison of particle size distributions (PSDs) is a fundamental task in industrial particle technology and process engineering, where particle size critically governs both process performance and product functionality. Across applications such as mineral processing [1, 2], mechanical processing [3, 4], separation systems [5, 6], additive manufacturing [7, 8], spray drying [9], agglomeration [10–13], filtration [14, 15], particle coating [16–18], combustion [19, 20],

*Corresponding author
Email address: Rahul.Mitra@mvtat.tu-freiberg.de (Rahul Mitra)

recycling [21–23], and food processing [24, 25], particle sizes serve as key indicators of process stability, efficiency, and product quality. For example, in a copper concentrator grinding circuit, the PSD of the ground ore is directly related to the degree of mineral liberation, thus influencing recovery efficiency and reagent consumption [26]. In lithium-ion battery recycling, PSD of the recovered cathode material affects the leaching kinetics and purity of lithium compounds [27]. Similarly, in the recycling of automotive components, the PSD of the crushed feed determines the selectivity and throughput of magnetic and eddy current separators [28, 29]. In additive manufacturing, powder PSD governs flowability, packing density [30, 31], and the microstructure of printed parts [32]. Across all these applications, even small changes in PSD can signal process drift, tool wear, or quality differences, making quantitative PSD comparison essential for process monitoring, control, and optimization.

Beyond these operational considerations, PSDs also underpin product functionality through the so-called "property function" introduced in [33]. This concept formalizes the relationship between the microscopic properties of the particle system and the macroscopic product properties by linking particle size, shape, morphology, and surface characteristics to functional outcomes such as flowability, reactivity, or optical behavior. This framework has been extended in [34], highlighting its significance in product engineering, where control of the PSD enables tailoring of particle characteristics to obtain the desired product properties. Thus, PSD is a fundamental design variable connecting process parameters to functional product behavior. Recognizing its fundamental importance, digital particle databases [35, 36] have been created that support benchmarking and can be accessed to retrieve particle size or shape distributions for subsequent comparison. The accurate and quantitative comparison of PSDs between different process conditions or batches is therefore indispensable for maintaining consistent product properties, ensuring specification compliance, and enabling data-driven product design.

To describe PSDs mathematically, several models are commonly used and are intended to represent the underlying probability density of particle sizes by analytical formulas. In most engineering applications, these are parametric distributions characterized by a certain number of model parameters (typically 2 or 3), which provide compact descriptors of a PSD in terms of a characteristic size and a spread. As summarized in [37], the most frequently used models include normal, lognormal, Rosin-Rammler-Sperling-Bennett (RRSB) (also commonly referred to as the Weibull distribution), and Gates-Gaudin-Schuhmann (GGS) distributions. The family of normal distributions represents symmetric PSDs, whereas the lognormal distribution suits growth processes such as agglomeration and spray drying, yielding the characteristic right-skewed tails. The RRSB distribution describes fragmented or milled materials and allows straightforward fitting to sieve data, whereas the GGS distribution captures the behavior of PSDs from mineral grinding or crushing operations and similar processes. Each of these parametric distributions provides a characteristic parameter useful for industrial process monitoring. However, such simplified approximations often fail to capture the multimodal features of particle systems. Beyond descriptive fitting of PSD models, the derivation of particle structure-property relationships is the fundamental reason why PSDs are predicted using mechanistic and data-driven models in the first place [38, 39]. Therefore, a quantitative comparison of predicted and measured PSDs is essential to assess whether a model adequately captures the particle population relevant for model validation, scale-up, and process optimization. For example, population-balance agglomeration models have been tested by directly comparing predicted distributions with experimental observations [40, 41]. During comparison, localized differences highlight size regions where the models under-predict or over-predict observed probabilities. Therefore, this validation step ensures that PSD-based parameters reliably represent the particle system, strengthening process understanding and product

optimization in particle technology.

Comparison of PSDs in practice often starts with qualitative tools such as visual overlays and difference-of-probability plots [42], which provide useful impressions but cannot assess whether observed differences are statistically significant. As a result, they do not support rigorous quantitative inference. Classical global tests such as the Pearson χ^2 [43–45] and Kolmogorov-Smirnov (KS) [46] tests evaluate overall similarities but compress all differences into a single scalar statistic. More recent distance-based measures, including the Jeffries-Matusita distance [47], total variation [48], maximum mean discrepancy (MMD) [49], Jensen-Shannon divergence [50], and the Wasserstein distance [51] offer improved sensitivity but still yield only a global measure of dissimilarity. Q-Q plots [52] are also commonly used to visually compare distributional agreement but do not provide localized information on where differences originate along the size axis (bins). These approaches cannot identify where along the particle size axis the distributions differ, which is often the most relevant information in engineering practice, where knowing the specific size ranges responsible for differences is essential for assessing process behavior and implementing corrective action.

Some localized methods have been proposed to address this need. In [53], a per-bin significance approach has been introduced that is based on normalized differences between observed and expected counts, offering an initial attempt at localized inference. However, its statistical validity is limited by the use of approximate variance models and by the absence of multiple-comparison control. In the work by Sharpe [54], residual-based and partitioning methods have been evaluated to locate sources of significance in χ^2 tests but they have concluded that these approaches lack formal control of the family-wise error rate (FWER), often leading to inflated false-positives in large datasets [55, 56]. In other words, when many bins are tested simultaneously, some bins may appear to differ purely by chance, even if the overall distributions are actually identical. Controlling the FWER ensures that the probability of incorrectly identifying any bin as significantly different remains within a predefined significance level, thereby distinguishing genuine process differences from random statistical noise. Although these studies represent important methodological advances, they remain insufficient for a detailed PSD comparison, where bin-wise correlations and multiplicity must be properly addressed.

Building upon these developments, in the present work we introduce a bin-resolved PSD comparison framework, termed the Difference Significance Profile (DSP), designed for particle-discrete PSD datasets. The DSP standardizes per-bin probability differences using a pooled-variance estimator and calibrates a global decision threshold using parametric bootstrap [57] under the null hypothesis that both PSDs originate from the same parent distribution. This allows simultaneous bin-wise inference with reduced false-positive detections. Previous applications of bootstrap resampling in particle technology and mineral processing [58–60] and extensive error analysis [61, 62] demonstrated the robustness of the method for uncertainty quantification in size and composition measurements, ranging from confidence interval (CI) estimation in single PSDs to error propagation in mineralogical characterization and separation performance evaluation. We extend this bootstrap-based concept from uncertainty estimation to inferential comparison between two PSDs, enabling statistically valid identification of bins where two distributions differ significantly. Therefore, DSP transforms qualitative histogram comparison into a quantitative and bin-resolved tool for simultaneous inference that directly supports process control, validation of product specifications, and data-driven decision-making in particle technology.

The paper is organized as follows. Section 2 introduces the statistical basis of the proposed DSP framework, including the common-bin representation of PSDs, the multinomial sampling model, local DSP statistics, and the bootstrap-based simultaneous calibration used for bin-wise

inference. Section 3 demonstrates the implementation of the framework on experimental spray-drying data and discusses its behavior for both different process conditions and repeat experiments under identical conditions. Section 4 summarizes the main findings and outlines possible directions for the future development of the framework.

2. The proposed framework: Difference Significance Profile

The proposed framework provides a statistically grounded and practically interpretable procedure for comparing two particle size distributions at the resolution of individual bins. Its aim is to quantify the observed bin-wise differences between two samples, standardize these differences relative to their estimated sampling uncertainty, and identify those bins whose differences are large enough to remain significant after simultaneous calibration across all bins. In this way, DSP combines local resolution with bootstrap-based simultaneous error control under the fitted multinomial model, so that marked bins represent size classes in which the observed differences are statistically significant within this model-based framework.

This section first presents the mathematical basis of the framework, including the common-bin statistical model, local standardized statistics, and simultaneous calibration based on bootstrap resampling. This formal development is then complemented with a practical engineering interpretation of the DSP framework, emphasizing how it should be read and used in particle-process analysis.

2.1. Observed data and common binning

Let two particle populations (e.g., obtained from two process conditions) be represented by the observed samples

$$x_1 = (x_{11}, \ldots, x_{1n_1}) \in \mathbb{R}_{>0}^{n_1}, \quad x_2 = (x_{21}, \ldots, x_{2n_2}) \in \mathbb{R}_{>0}^{n_2}, \tag{1}$$

where $x_{jr} \in \mathbb{R}_{>0} = (0, \infty)$ denotes the observed size of particle $r \in \{1, \ldots, n_j\}$ in sample $j \in \{1, 2\}$, and $n_j \in \mathbb{N}$ is the number of particles observed in sample j. In practice, n_j is determined by the measurement technique and sampling strategy, e.g., the number of particles detected in dynamic imaging.

A PSD is typically represented in binned form. For $j \in \{1, 2\}$ and an observed sample $x_j \in \mathbb{R}^{n_j}$, a binning is defined by a set of bin edges $e_0^{(j)}, \ldots, e_{k_j}^{(j)} \in \mathbb{R}$ with

$$e_0^{(j)} < e_1^{(j)} < \ldots < e_{k_j}^{(j)}, \tag{2}$$

for $k_j \in \mathbb{N} = \{1, 2, \ldots\}$, which induces a partition of the real axis into k_j bins. The observed count $n_{ji} \in \mathbb{N}_0$ in bin i is given by

$$n_{ji} = \#\left\{r \in \{1, \ldots, n_j\} : x_{jr} \in (e_{i-1}^{(j)}, e_i^{(j)}]\right\} \in \mathbb{N}_0, \quad i \in \{1, \ldots, k_j\}, \tag{3}$$

where # denotes the cardinality of a set. The height of bin i is then given by the relative frequency or the empirical bin probability $\hat{p}_{ji} \in [0, 1]$,

$$\hat{p}_{ji} = \frac{n_{ji}}{n_j} \in [0, 1], \tag{4}$$

where $\sum_{i=1}^{k_j} \hat{p}_{ji} = 1$.

A comparison of the empirical probabilities (bin heights) $\{\hat{p}_{1i}\}$ and $\{\hat{p}_{2i}\}$ of two histograms with different bin edges $\{e_i^{(1)}\} \neq \{e_i^{(2)}\}$ is misleading. In that case, any bin-wise difference misrepresents the underlying distributional differences due to artifacts arising from the chosen binning. Therefore, a prerequisite for any bin-wise comparison is that both samples be evaluated on a common set of bins. To this end, a joint binning is introduced with bin edges

$$e_0 < e_1 < \cdots < e_k, \tag{5}$$

for $k \in \mathbb{N}$, which is then applied to both samples.

If there are no predefined bins, a data-adaptive choice of the bin edges $e_0, e_1, \ldots, e_k$ for some $k \in \mathbb{N}$ can be obtained from the pooled sample x, which is given by

$$x = (x_{11}, \ldots, x_{1n_1}, x_{21}, \ldots, x_{2n_2}) = (x_1, \ldots, x_n) \in \mathbb{R}^n_{>0}, \tag{6}$$

where $n = n_1 + n_2 \in \mathbb{N}$ is the total sample size and we assume that there exists $r, s \in \{1, \ldots, n\}$, $r \neq s$ such that $x_r \neq x_s$. Here, the Freedman-Diaconis rule [63] can be used for a constant bin width $h(x) = e_i - e_{i-1}$ for all $i \in \{1, \ldots, k\}$, given by

$$h(x) = \frac{2\,\mathrm{IQR}(x)}{n^{1/3}} \in \mathbb{R}_{\geq 0}, \tag{7}$$

where $\mathrm{IQR}(x) = Q_{0.75}(x) - Q_{0.25}(x)$ is the interquartile range and $Q_q(x)$ denotes the q quantile for $q \in (0, 1)$. The resulting number of bins $k > 0$ is then given by

$$k = \left\lceil \frac{x_{\max} - x_{\min}}{h(x)} \right\rceil, \tag{8}$$

with $x_{\min}$ and $x_{\max}$ being the smallest and largest component of x, respectively, and $\lceil \cdot \rceil \colon \mathbb{R} \to \mathbb{Z} = \{\ldots, -1, 0, 1, \ldots\}$ denotes rounding to the smallest integer greater than or equal to the argument. The corresponding bin edges are then defined as

$$e_i = e_0 + i\,h(x), \tag{9}$$

for each $i \in \{1, \ldots, k\}$, where the first bin $[e_0, e_1]$ is a closed interval (with $e_0 = x_{\min}$), and the remaining bins are half-open intervals $(e_{i-1}, e_i]$ for $i \in \{2, \ldots, k\}$ (with $e_k = x_{\min} + kh \geq x_{\max}$). This ensures that all data points belonging to the interval $[x_{\min}, x_{\max}]$ fall into some bin and that both samples x_1 and x_2 are evaluated on an identical set of bins, allowing direct bin-wise comparison. To keep the notation simple, the bins (after common binning) are denoted by

$$\mathrm{E}_i = \begin{cases} [e_{i-1}, e_i], & i = 1, \\ (e_{i-1}, e_i], & i > 1. \end{cases} \tag{10}$$

If $\mathrm{IQR}(x)$ is close to zero, then the bin width $h(x)$ also approaches zero, which implies an extremely large number of bins k. To avoid this degenerate situation, a maximum number of bins can be heuristically chosen to ensure a meaningful partition.

Depending on the data, other common binning techniques can also be used, such as Scott's rule [64], Sturges' rule [65], or more advanced binning strategies [66]. However, in the present paper, the Freedman-Diaconis rule is used as a data-adaptive starting point that determines the common bin width from the pooled sample size and spread through the interquartile range.

In many practical settings, the bins are prescribed by the measurement technique itself, for example, in the case of dynamic image analysis, and may therefore have unequal or logarithmically spaced widths. DSP directly accommodates such binning schemes because its statistical formulation is based on the counts and probability masses within the bins. The only essential requirement is that both samples be represented using the same set of bin edges.

2.2. Statistical model for the common-bin counts

This section introduces the mathematical foundation behind DSP. For this, let $\mathrm{E}_1, \ldots, \mathrm{E}_k,\ k \in \mathbb{N}$ be some fixed binning. For each sample $j \in \{1, 2\}$ with sample size $n_j \in \mathbb{N}$, assume that the measured particle sizes follow independent and identically distributed random variables $X_{j1}, \ldots, X_{jn_j}$ taking values in $\bigcup_{i=1}^{k} \mathrm{E}_i$. As in Eq. (4), define $p_{ji} \in [0, 1]$ as the probability of X_{j1} falling in bin i, i.e.,

$$p_{ji} = \mathbb{P}(X_{jr} \in \mathrm{E}_i). \tag{11}$$

The corresponding random bin count N_{ji} is given by

$$N_{ji} = \#\Big\{r \in \{1, \ldots, n_j\} : X_{jr} \in \mathrm{E}_i\Big\}, \tag{12}$$

and the random count vector $\mathrm{N}_j = (N_{j1}, \ldots, N_{jk}) \in \mathbb{N}_0^k$ is multinomially distributed with the parameter vector $\mathrm{p}_j = (p_{j1}, \ldots, p_{jk}) \in [0, 1]^k$ [67], i.e.,

$$\mathrm{N}_j \sim \mathrm{Mult}(n_j; \mathrm{p}_j). \tag{13}$$

The model implies inter-bin dependence and negative covariance between the random bin counts $(N_{j1}, \ldots, N_{jk})$, with covariance and variance given by

$$\mathrm{Cov}(N_{ji}, N_{j\ell}) = -n_j p_{ji} p_{j\ell} \quad (i \neq \ell), \qquad \mathrm{Var}(N_{ji}) = n_j p_{ji}(1 - p_{ji}). \tag{14}$$

This multinomial sampling view is particularly appropriate for number-weighted PSDs obtained from particle-wise measurement techniques, such as image-based analysis, optical particle counters, or in-line imaging probes. Under the working assumption that the detected particles can be treated as independent and identically distributed draws from the same underlying particle population, assigning each particle to a bin corresponds to a categorical outcome with mutually exclusive alternatives. The multinomial distribution is therefore the natural probabilistic model for this counting process under that working assumption. If, however, there are strong within-sample dependencies or additional sources of variability induced by the measurement process, then the data may exhibit over-dispersion relative to the multinomial model. In that case, one may mitigate this by using repeated samples, by introducing an effective sample size, or by adopting an over-dispersed extension such as a Dirichlet-multinomial distribution [68, 69]. Nevertheless, for number-weighted PSDs, the multinomial model remains a natural working model that captures the compositional dependence induced by binning.

The goal of DSP is to test the null hypothesis $H_0 : \mathrm{p}_1 = \mathrm{p}_2$ for the two observed histograms, where each histogram is modeled as an independent multinomial sample on the same binning. DSP evaluates the bin-wise difference of histograms relative to the bin-wise uncertainty under the null hypothesis H_0 of equal bin probabilities across both samples.

Note that under the null hypothesis $H_0 : \mathrm{p}_1 = \mathrm{p}_2$, which means in particular that $p_{1i} = p_{2i}$ for each $i \in \{1, \ldots, k\}$, the standard error SE_i of the estimator $(N_{1i}/n_1) - (N_{2i}/n_2)$ of the "true" bin-wise difference $p_{1i} - p_{2i} = 0$ is given by

$$\mathrm{SE}_i = \sqrt{p_i(1 - p_i)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} \in \mathbb{R}_{\geq 0}. \tag{15}$$

Here, $p_i = p_{1i} = p_{2i}$ denotes the common "true" probability that a randomly selected particle falls into bin i under H_0. The standard error directly quantifies the uncertainty, i.e., the typical range

of variation one would expect in a bin-wise difference due to random sampling, even when the two distributions are truly identical. The uncertainty decreases as more particles are sampled and is modulated by how strongly populated a bin is. In practice, this means that comparing small sample sizes or sparse bins naturally produces lower statistical power.

2.3. Observed local DSP statistics

Given the observed samples x_1 and x_2 and the fixed common bins $\mathrm{E}_1, \ldots, \mathrm{E}_k$, the observed bin counts are

$$n_{ji} = \#\Big\{r \in \{1, \ldots, n_j\} : x_{jr} \in \mathrm{E}_i\Big\}. \tag{16}$$

In other words, n_{ji} is the number of particles in sample j whose measured sizes belong to the i-th bin after fixing the common bin intervals E_i. Similarly to the definition introduced in Eq. (4), the corresponding empirical bin probability within E_i is given by

$$\hat{p}_{ji} = \frac{n_{ji}}{n_j} \in [0, 1], \tag{17}$$

where $\sum_{i=1}^{k} \hat{p}_{ji} = 1$. The empirical bin probability $\hat{p}_{ji}$ can be regarded as a realization of the consistent estimator N_{ji}/n_j of the underlying probability p_{ji} (introduced in Eq. (11)).

For the observed data, the empirical bin-wise difference is defined as

$$\Delta_i = \hat{p}_{1i} - \hat{p}_{2i} \in [-1, 1], \tag{18}$$

for each $i \in \{1, \ldots, k\}$, where Δ_i can be regarded as a realization of the consistent estimator $(N_{1i}/n_1) - (N_{2i}/n_2)$ of the "true" bin-wise difference $p_{1i} - p_{2i}$. For example, $\Delta_i > 0$ indicates enrichment of bin i in sample 1 relative to sample 2, while $\Delta_i < 0$ indicates depletion.

Under the null hypothesis H_0, both samples are assumed to share the same underlying bin probability p_i in bin i. It is natural to estimate p_i by pooling the observed counts from the two samples. Therefore, a realization $\hat{p}_i$ of the pooled estimator of p_i is given by

$$\hat{p}_i = \frac{\sum_{j=1}^{2} n_{ji}}{\sum_{j=1}^{2} n_j} = \frac{n_{1i} + n_{2i}}{n_1 + n_2} \in [0, 1]. \tag{19}$$

Thus, $\hat{p}_i$ is simply the empirical probability of all observed particles, from both samples combined, that fall into bin i. In other words, one may imagine merging the two samples into a single pooled sample of size $n_1 + n_2$ and then calculating the fraction of particles in that pooled sample that belong to bin i. From a statistical point of view, $\hat{p}_i$ is also a realization of the maximum-likelihood estimator [70] of the "true" common bin probability p_i. Intuitively, it is the best-supported estimate of the probability of bin i when one assumes that there is no true distributional difference between the two samples in that bin.

This estimator is appropriate under H_0 because, when both samples truly have the same bin probability p_i, combining them increases the amount of information available for estimating that common probability. As a result, $\hat{p}_i$ is typically more stable than the realization of $\hat{p}_{1i}$ and $\hat{p}_{2i}$ of separate estimators considered individually. Equivalently, $\hat{p}_i$ can be interpreted as a weighted average of the two sample-specific bin frequencies,

$$\hat{p}_i = \frac{n_1}{n_1 + n_2}\hat{p}_{1i} + \frac{n_2}{n_1 + n_2}\hat{p}_{2i}, \tag{20}$$

so that the sample with the larger number of observed particles contributes more strongly to the pooled estimate.

Since the "true" standard error SE_i under H_0, from Eq. (15), is not directly available from the data, DSP uses the consistent estimator whose realizations $\widehat{\mathrm{SE}}_i$ are given by

$$\widehat{\mathrm{SE}}_i = \sqrt{\hat{p}_i(1-\hat{p}_i)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}. \tag{21}$$

DSP computes the bin-wise standardized (dimensionless) statistic (the Z-score) [71], denoted by $Z_i \in \mathbb{R}$, as

$$Z_i = \begin{cases} \dfrac{\Delta_i}{\widehat{\mathrm{SE}}_i}, & \widehat{\mathrm{SE}}_i > 0, \\ 0, & \widehat{\mathrm{SE}}_i = 0. \end{cases} \tag{22}$$

Thus, Z_i standardizes the observed bin-wise difference Δ_i by its estimated standard error $\widehat{\mathrm{SE}}_i$. Intuitively, Z_i measures how large the observed bin-wise difference is relative to the sampling uncertainty expected under the null hypothesis.

It is important to note that $\widehat{\mathrm{SE}}_i = 0$ whenever the pooled estimator is degenerate, i.e., when $\hat{p}_i \in \{0, 1\}$. The practically relevant case is usually $\hat{p}_i = 0$, which means that bin i is empty in both observed samples. In such a case, the ratio in Eq. (22) is not defined. These bins are therefore treated as non-informative and will not be marked as significant. This convention is consistent with the interpretation that a bin with no pooled variation provides no evidence of a local difference. Therefore, Z_i is explicitly set to zero whenever $\widehat{\mathrm{SE}}_i = 0$.

2.4. Global difference through Pearson's χ^2 statistic

Complementing the bin-wise DSP statistics, the overall difference between the two binned PSDs is summarized by Pearson's χ^2 statistic. In the present framework, this quantity is used primarily as a global difference measure under the null hypothesis H_0 that both samples share the same bin-probability vector in all bins. Consequently, the expected count E_{ji} in bin i for sample $j \in \{1, 2\}$ is given by

$$E_{ji} = n_j p_i. \tag{23}$$

Thus, E_{ji} represents the count that would be expected in bin i for sample j if both samples were generated from the same common PSD. However, in practice, the "true" bin probability p_i is unknown and must be estimated from the observed data. As in Eq. (19), the pooled estimate $\hat{p}_i$ is used, which yields the estimated expected count

$$\hat{E}_{ji} = n_j \hat{p}_i. \tag{24}$$

Hence, $\hat{E}_{ji}$ is the expected count in bin i for sample j under the fitted model obtained from the pooled data. The observed Pearson statistic $\chi^2 \in [0, \infty)$ is then defined as

$$\chi^2 = \sum_{i=1}^{k} \sum_{\substack{j=1 \\ \hat{E}_{ji} \neq 0}}^{2} \frac{(n_{ji} - \hat{E}_{ji})^2}{\hat{E}_{ji}}. \tag{25}$$

Note that bins with $\hat{E}_{ji} = 0$ are ignored in the summation considered in Eq. (25). This can occur only when both bin counts are zero, in which case the bin contains no information about the difference between the two samples and contributes nothing to the statistic.

The Pearson statistic provides a single scalar measure of the global difference across all bins. If desired, it may be decomposed into bin-wise contributions,

$$\chi_i^2 = \sum_{j=1}^{2} \frac{(n_{ji} - \hat{E}_{ji})^2}{\hat{E}_{ji}}, \tag{26}$$

so that

$$\chi^2 = \sum_{i=1}^{k} \chi_i^2. \tag{27}$$

These quantities indicate how strongly individual bins contribute to the total Pearson difference. However, they are not the quantities on which DSP bases its local interpretation.

The reason is twofold. First, each χ_i^2 is non-negative and therefore does not encode the sign of the difference. A large value of χ_i^2 indicates that the observed count in bin i differs from its null hypothesis expectation, but does not indicate whether sample 1 is enriched or depleted relative to sample 2 in that bin. In contrast, the DSP statistic Z_i considered in Eq. (22) retains the sign and therefore directly indicates the direction of the observed local shift. The second reason for basing DSP on local statistics Z_i rather than on Pearson contributions χ_i^2 is the interpretability on the probability scale. The Z_i directly quantifies how the bin probabilities differ between the two samples relative to their standard error under H_0. Thus, the sign indicates whether sample 1 is enriched or depleted relative to sample 2, and the magnitude indicates how large that probability difference is compared to ordinary random variation. By contrast, the bin-wise Pearson contribution χ_i^2 is not directly interpretable as a difference of probabilities. It does not indicate which sample has the larger bin probability. For this reason, χ_i^2 is useful for decomposing the global difference, while Z_i is the more natural statistic for local interpretation and bin-wise inference.

In practice, DSP therefore uses the Pearson χ^2 statistic as a global difference measure, while the local statistics Z_i provide bin-resolved information about where the observed differences are large relative to their uncertainty.

2.5. Simultaneous calibration by parametric bootstrap and decision rule

A key practical challenge in bin-wise PSD comparison is that many bins are evaluated simultaneously. In process monitoring or model validation, this corresponds to asking many localized questions at once (one per bin). Without a simultaneous calibration, one would expect occasional large $|Z_i|$ values purely due to random sampling, especially for fine binning. Since $Z_1, \ldots, Z_k$ are correlated across bins under the multinomial model, DSP determines a single simultaneous threshold that is intended to provide approximate family-wise error control under the fitted multinomial model, i.e., approximate control of the probability of marking at least one bin as significant under H_0. This is achieved using a parametric bootstrap [57, 72].

Under H_0, the pooled empirical probability vector $\hat{\mathrm{p}} = (\hat{p}_1, \ldots, \hat{p}_k) \in \mathcal{P}_k = [0, 1]^k$ serves as the fitted multinomial model. Conditional on the chosen common bins and the observed sample sizes n_1 and n_2, bootstrap count vectors are generated

$$\mathrm{N}_1^{*(b)} \sim \mathrm{Mult}(n_1; \hat{\mathrm{p}}), \quad \mathrm{N}_2^{*(b)} \sim \mathrm{Mult}(n_2; \hat{\mathrm{p}}), \quad b \in \{1, \ldots, B\}, \tag{28}$$

for some $B \in \mathbb{N}$. This multinomial resampling step reproduces the dependence induced by distributing a fixed number of particles over a common set of mutually exclusive bins. In particular,

it preserves the negative inter-bin dependence that is characteristic of binned number-weighted PSDs.

For each bootstrap iteration $b \in \{1, \dots, B\}$, the bootstrap analogs of the empirical probabilities are computed,

$$\hat{p}_{ji}^{*(b)} = \frac{N_{ji}^{*(b)}}{n_j}, \quad \hat{p}_i^{*(b)} = \frac{N_{1i}^{*(b)} + N_{2i}^{*(b)}}{n_1 + n_2}, \tag{29}$$

together with the corresponding bootstrap difference profile,

$$\Delta_i^{*(b)} = \hat{p}_{1i}^{*(b)} - \hat{p}_{2i}^{*(b)}, \tag{30}$$

the local standard error estimates,

$$\widehat{\mathrm{SE}}_i^{*(b)} = \sqrt{\hat{p}_i^{*(b)}(1 - \hat{p}_i^{*(b)})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}, \tag{31}$$

and the corresponding standardized differences,

$$Z_i^{*(b)} = \begin{cases} \dfrac{\Delta_i^{*(b)}}{\widehat{\mathrm{SE}}_i^{*(b)}}, & \widehat{\mathrm{SE}}_i^{*(b)} > 0, \\ 0, & \widehat{\mathrm{SE}}_i^{*(b)} = 0. \end{cases} \tag{32}$$

Each bootstrap replicate is then summarized by the maximum absolute standardized difference across all bins,

$$M^{*(b)} = \max_{i \in \{1,\dots,k\}} |Z_i^{*(b)}|. \tag{33}$$

Thus, the bootstrap sample $M^{*(1)}, \dots, M^{*(B)}$ approximates the distribution of the largest bin-wise fluctuation that may occur when both samples arise from the same underlying PSD.

Based on these bootstrap maxima, the empirical simultaneous threshold is defined as the $(1 - \alpha)$-quantile, given by

$$c_{1-\alpha}(B) = Q_{1-\alpha}(M^{*(1)}, \dots, M^{*(B)}). \tag{34}$$

After calibrating the simultaneous threshold using the parametric bootstrap, each observed Z_i is compared with the bootstrap threshold $c_{1-\alpha}(B)$, where bin i is marked significant whenever

$$|Z_i| > c_{1-\alpha}(B). \tag{35}$$

In addition to the bootstrap maxima used for the simultaneous calibration of the local statistics Z_i, the same parametric bootstrap can also be used to assess the global Pearson discrepancy. For each bootstrap iteration $b \in \{1, \dots, B\}$, the bootstrap analog of Pearson's χ^2 statistic is computed,

$$\chi^{2*(b)} = \sum_{i=1}^{k} \sum_{j=1}^{2} \frac{(N_{ji}^{*(b)} - \hat{E}_{ji}^{*(b)})^2}{\hat{E}_{ji}^{*(b)}}, \tag{36}$$

where

$$\hat{E}_{ji}^{*(b)} = n_j \hat{p}_i^{*(b)}. \tag{37}$$

Therefore, the bootstrap sample $\chi^{2*(1)}, \ldots, \chi^{2*(B)}$ approximates the distribution of Pearson's χ^2 statistic under the fitted multinomial model. This allows the bootstrap p-value to be computed

$$p_{\text{boot}} = \frac{1 + \sum_{b=1}^{B} \mathbf{1}\left(\chi^{2*(b)} \geq \chi^2\right)}{B + 1}, \tag{38}$$

where χ^2 denotes the calculated Pearson statistic from Eq. (25) and $\mathbf{1}(\cdot)$ is the indicator function, defined by

$$\mathbf{1}(A) = \begin{cases} 1, & \text{if the event } A \text{ is true,} \\ 0, & \text{if the event } A \text{ is false.} \end{cases} \tag{39}$$

Therefore, $\mathbf{1}\left(\chi^{2*(b)} \geq \chi^2\right)$ takes the value 1 if $\chi^{2*(b)}$ is at least as large as χ^2, and 0 otherwise. Thus, p_{boot} provides a model-based global significance assessment for the overall difference between the two binned PSDs, complementing the bin-resolved DSP decision rule from Eq. (35).

2.6. Interpretation of the simultaneous threshold and marked bins

To formalize the interpretation of the threshold considered in Section 2.5, let M denote the random maximum absolute standardized Z-score under H_0, which is given by

$$M = \max_{i \in \{1,\ldots,k\}} |Z_i|. \tag{40}$$

The bootstrap maxima $M^{*(1)}, \ldots, M^{*(B)}$ are used to approximate the distribution of M. Therefore, for some significance level $\alpha \in (0, 1)$, the simultaneous threshold is the $(1 - \alpha)$-quantile of the distribution of M, given by

$$t_{1-\alpha} = Q_{1-\alpha}(M). \tag{41}$$

Thus, $c_{1-\alpha}(B)$ from Eq. (34) is the bootstrap estimate of $t_{1-\alpha}$, based on B resamples.

Now, by construction,

$$\mathbb{P}_{H_0}(M > t_{1-\alpha}) = \mathbb{P}_{H_0}(\text{there is an } i \in \{1, \ldots, k\} : |Z_i| > t_{1-\alpha}) \leq \alpha. \tag{42}$$

Therefore, Eq. (42) states that, under the null hypothesis H_0, the probability that there exists at least one bin whose standardized difference exceeds the ideal simultaneous threshold $t_{1-\alpha}$ is at most α. If the distribution of M is continuous at $t_{1-\alpha}$, then the left-hand side equals α. In practice, DSP replaces $t_{1-\alpha}$ by its parametric-bootstrap estimate $c_{1-\alpha}$. Consequently, the implemented procedure is intended to provide approximate simultaneous control of the probability of obtaining one or more false-positives anywhere in the collection of bins under the fitted multinomial model.

At this point, it is important to distinguish clearly between the local and global aspects of DSP. The quantities Δ_i, $\widehat{\text{SE}}_i$, and Z_i are all computed separately for each bin and therefore provide a bin-resolved description of where the two PSDs differ and in which direction the local difference points. In this descriptive sense, DSP is local. However, the statistical decision is not made bin-by-bin in isolation. Instead, the decision threshold is calibrated by the distribution of the maximum statistic $M^{*(1)}, \ldots, M^{*(B)}$, i.e., through the joint behavior of the entire vector $(Z_1, \ldots, Z_k)$. Hence, DSP combines bin-resolved local differences with global simultaneous inference. It reports a per-bin difference profile but evaluates those bin-wise differences using one common threshold that is calibrated from the whole histogram. This is also why a bin may show a moderately large value of $|Z_i|$ when viewed in isolation and not be marked as significant, because the observed

difference is not large enough after accounting for the fact that many correlated bins are examined simultaneously.

This simultaneous interpretation must be distinguished carefully from a point-wise one. For any bin i,

$$\mathbb{P}_{H_0}(|Z_i| > t_{1-\alpha}) \leq \mathbb{P}_{H_0}\Big(\max_{1\leq \ell \leq k} |Z_\ell| > t_{1-\alpha}\Big) \leq \alpha. \tag{43}$$

Eq. (43) shows that exceedance of the threshold in one specific bin is only a special case of exceedance somewhere among all bins. Therefore, $t_{1-\alpha}$ is not a point-wise $100(1-\alpha)\%$ cutoff for each individual bin. In particular, one should not interpret a marked bin as a bin that is "different with probability $100(1-\alpha)\%$" or that each bin is assessed by an ordinary marginal confidence interval. Instead, the correct interpretation is that, for the ideal threshold $t_{1-\alpha}$, the overall probability of obtaining one or more false-positives anywhere in the entire set of bins is bounded by α. The bootstrap threshold $c_{1-\alpha}$ is intended to approximate this simultaneous control under the fitted multinomial model.

This simultaneous calibration is especially important in particle technology applications, because false-positive findings may lead to unnecessary process interventions or incorrect scientific conclusions. For example, one might otherwise adjust operating conditions, reject a batch, or conclude that a predictive model fails to reproduce the PSD, even though the apparent difference is within the limits of random sampling variability. By calibrating against the maximum standardized Z_i over all bins, DSP is intended to ensure that marked bins correspond to local differences that remain statistically significant under the fitted multinomial model after accounting for multiplicity.

The marked bins are collected in the set $\mathcal{S}_\alpha$, given by

$$\mathcal{S}_\alpha = \Big\{ i \in \{1, \ldots, k\} : |Z_i| > c_{1-\alpha} \Big\}. \tag{44}$$

If this set is not empty, i.e. $\mathcal{S}_\alpha \neq \varnothing$, then the null hypothesis H_0 is rejected within the bootstrap-calibrated multinomial model. Under this model assumption, this means that the observed PSD difference profile contains at least one local difference that is too large to be explained by finite-sample variation alone at the chosen simultaneous level α. Equivalently, at least one bin shows a difference that remains statistically significant after accounting for the fact that all bins are evaluated simultaneously. If $\mathcal{S}_\alpha = \varnothing$, then H_0 is not rejected. This does not prove that the two PSDs are identical in every bin. Rather, it means that none of the observed bin-wise differences is large enough, relative to its uncertainty under H_0, to exceed the simultaneous threshold in the present dataset.

This distinction is important. The simultaneous threshold is intended to provide approximate control of the probability of falsely marking at least one bin under the null hypothesis H_0. It does not assign a posterior probability to any individual bin being "the cause" of the difference, nor does it imply that each fixed bin exceeds the threshold with probability α. Instead, DSP should be interpreted as a framework that marks the bins whose observed differences are large enough to remain significant after accounting for the fact that the entire family of bins is tested at once. In this sense, the set $\mathcal{S}_\alpha$ identifies the size classes in which local PSD differences are statistically significant at the chosen simultaneous level α, under the fitted multinomial model. At the same time, smaller differences may still be present, but not significant at the selected simultaneous level $\alpha \in (0, 1)$.

Finally, since the observed difference profile is compositional,

$$\sum_{i=1}^{k} \Delta_i = 0, \tag{45}$$

a significant increase in one region of the PSD must be accompanied by compensating decreases elsewhere. Some of these compensating shifts may remain below the simultaneous threshold. Consequently, even when only one bin is marked, the underlying physical change should generally be understood as a redistribution of probability mass across the PSD, with bin i representing the dominant statistically resolved location of this redistribution.

2.7. Practical interpretation of the DSP framework

In practical terms, DSP is meant to answer two simple questions: (1) where do two PSDs differ, and (2) are those observed differences large enough to be more than ordinary random fluctuation? The framework begins by placing both PSDs on the same common bins, as described in Section 2.1. This step is essential because only then do the bins represent the same size classes in both samples.

Once the common bins are fixed, DSP computes the bin-wise difference Δ_i in Eq. (18). This quantity indicates how the relative amount of particles in bin i differs between the two samples. A positive value means that sample 1 contains a larger fraction of particles in that size class than sample 2, while a negative value means that sample 1 contains a smaller fraction. Therefore, the sign of Δ_i already shows the direction of the local change along the particle-size axis.

However, not every observed difference should be interpreted as meaningful. Even when two underlying PSDs are truly the same, the measured bin heights will not match perfectly because only a finite number of particles are observed. For this reason, DSP also computes the estimated standard error $\widehat{\mathrm{SE}}_i$ and the standardized statistic Z_i in Eq. (22). The estimated standard error describes how much random variation would normally be expected in bin i due to finite sampling alone. The statistic Z_i then compares the observed difference Δ_i with the expected variation. In this sense, Z_i can be understood as a signal-to-noise measure that shows whether the observed local difference is small enough to be explained by ordinary sampling variation, or large enough to be marked as significant.

Since all bins are examined at the same time, DSP does not check each bin using an ordinary point-wise cutoff. If many bins are checked separately, some of them may appear unusual purely by chance, even when there is no real difference between the PSDs. To avoid this, DSP calibrates a single simultaneous threshold $c_{1-\alpha}$ by bootstrap resampling, as described in Section 2.5. Under the fitted model assumption, this threshold is chosen so that the probability of falsely marking at least one bin anywhere in the full PSD is approximately controlled at level α. Thus, DSP protects against false-positives that would otherwise arise from inspecting many bins at once.

The final DSP result is the set of marked bins, i.e., the bins for which $|Z_i| > c_{1-\alpha}$. These are the size classes in which the observed differences are too large to be plausibly explained by finite-sample variation alone under the fitted model. If no bins are marked, this does not prove that the PSDs are identical. This only means that in the available data, no local difference is large enough to exceed the simultaneous threshold. If one or more bins are marked, then DSP localizes the statistically significant differences to those bins.

In general, DSP should be considered a practical statistical tool for PSD comparison. The key quantities Δ_i, $\widehat{\mathrm{SE}}_i$, Z_i, and $c_{1-\alpha}$ describe where the probability mass is redistributed across the PSD, how uncertain the observed redistribution is given the available sample size, and which local differences are large enough to remain significant after accounting for the fact that many bins are examined simultaneously. In this sense, DSP shows which size classes differ, in which direction they differ, and which parts of the PSD therefore require attention in process analysis. For example, it can reveal whether a process produces excess fines, an underrepresented modal

range, or an overestimated coarse tail. However, DSP does not determine the unique process change required to make two PSDs similar again. Rather, it provides statistically calibrated local difference information that can be combined with process knowledge and mechanistic models to guide such a control. This combination of local resolution, simultaneous error control, and direct interpretability makes DSP especially useful for process diagnostics and batch comparability studies in powder and particle technology.

3. Implementation of DSP framework on experimental data

To demonstrate the practical relevance of the proposed DSP framework under realistic experimental conditions, we consider a case of in-line PSD monitoring during spray drying, where we compare two number-weighted particle size distributions measured at different drying temperatures while holding all other operating conditions constant. This isolates the effect of drying temperature on particle formation. In addition, we analyze a repeat spray drying run under nominally identical process settings as a negative-control comparison, for which we expect no intentional process-induced difference. Finally, we compare the bin-wise DSP results from this repeat spray drying comparison with existing frameworks [53, 54] to provide a practical check of robustness against noise-driven detections arising from sampling variability and measurement uncertainty. To illustrate how engineering conclusions depend on the chosen simultaneous threshold, we report additional DSP evaluations for the 90% and 99% simultaneous levels along with the main analysis at 95%.

In each analysis, we first overlay the PSDs to provide a qualitative context and then apply DSP to obtain bin-resolved differences Δ_i, standardized differences Z_i, and the simultaneous decision threshold $c_{1-\alpha}$ used to identify significant bins. We use a pooled data-adaptive binning, as described in Section 2.1, to obtain a common partition of the size range. We perform simultaneous threshold calibration using parametric bootstrap with $B = 40000$ iterations. This choice provides numerically stable threshold estimates so that the significant-bin classification is not unduly affected by variability in the bootstrap calibration. We carry out all computations in Python 3.12.4 on a standard desktop system equipped with a 13th Gen Intel(R) Core(TM) i5-1334U processor and 16 GB RAM, resulting in a typical computation time of approximately 5 s for $k = 75$ bins and $B = 40000$ bootstrap iterations. We provide supporting results on the induced inter-bin dependence and on the bootstrap distributions used for calibration in Appendix A.

We obtain all number-weighted size distributions from in-line 2D images of the dried powder, captured using a SOPAT PL camera probe (SOPAT GmbH, Berlin, Germany). The image processing pipeline is discussed in detail in [73]. The goal is to evaluate how the drying temperature influences the resulting PSDs while all feed suspension and atomization parameters are held constant. We compare two drying conditions: 120 °C (T1) and 145 °C (T2), representing moderate and elevated thermal loads in the drying process, respectively. We then analyze a third repeat run at 120 °C ($T1_{repeat}$) to assess the behavior of the framework under nominally identical process conditions and, in particular, its robustness against noise-driven detections in a negative-control setting.

3.1. Comparison of particles produced at different spray drying temperatures

We obtain the measured number-weighted PSDs (Fig. 1a) from $n_1 = 2881$ and $n_2 = 2658$ individual particles. Both distributions exhibit a positively skewed shape, typical of spray-dried powders, dominated by a primary mode between 3 µm and 5 µm. The higher temperature case

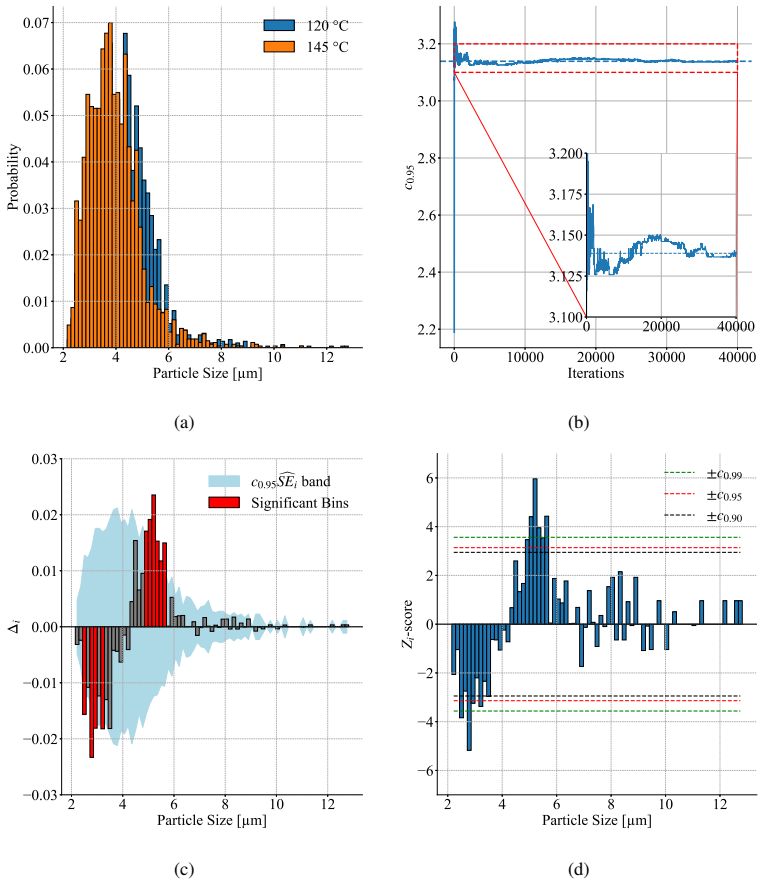


Figure 1: (a) Overlay of PSDs, (b) convergence of the simultaneous threshold $c_{0.95}$, (c) Δ_i with significant bins (red highlighted) outside the simultaneous reference band $\pm c_{0.95}\,\widehat{\mathrm{SE}}_i$, and (d) standardized Z_i for the spray-dried alumina samples at 120 °C (T1) and 145 °C (T2).

(T2) shows a distinct leftward shift of the mode and a narrower spread compared to T1, which is consistent with faster solidification and reduced droplet coalescence under the higher thermal load. Pooled binning using the Freedman–Diaconis rule yields $k = 75$ bins covering the size range of 2 µm to 13 µm. The global Pearson statistic $\chi^2 = 257$ with a bootstrap-calibrated $p_{\mathrm{boot}} \ll 0.001$ provides very strong evidence, under the fitted multinomial model, that the two drying conditions are associated with different PSDs at the global level. The bootstrap convergence plot (Fig. 1b) shows the simultaneous calibration threshold stabilizing at $c_{0.95} = 3.14$, demonstrating the numerical stability and reproducibility of the bootstrap-based correction.

The bin-wise differences Δ_i (Fig. 1c) show the main redistribution occurring in the 2 µm to 6 µm region. The maximum positive difference, $\Delta_i \approx 0.024$, occurs near 5.2 µm, where the 120 °C condition (T1) shows enrichment of these particles. In contrast, the largest negative difference, $\Delta_i \approx -0.021$, appears around 2.8 µm, where the 145 °C system produces more fines. About 13.3% of the bins exceed the simultaneously calibrated threshold at the 95% level and are marked in red. Here again, the simultaneous calibrated threshold band $\pm c_{1-\alpha}\widehat{\mathrm{SE}}_i$ should be interpreted

as a bin-wise quantity rather than as a continuous function of size. Significant differences are concentrated near the peak of the PSD, where most particles are found. This indicates that the process mainly redistributes particles among neighboring size classes in the dominant size range, rather than shifting the entire PSD uniformly toward finer or coarser sizes.

The standardized Z_i-scores (Fig. 1d) reinforce this conclusion. The largest differences, $Z_i \approx 5.9$ (at $\approx 5\,\mu m$) and $Z_i \approx -5.2$ (at $\approx 3\,\mu m$), both exceed the threshold $c_{0.95} = 3.14$ by a wide margin, indicating strong localized differences relative to the sampling variation expected under the fitted multinomial model. This abrupt alternation of positive and negative lobes signifies a reorganization of the particles from coarser to finer sizes under increased thermal load.

Physically, the observed shift toward finer particles may be explained by two plausible hypotheses related to the temperature-dependent drying kinetics. The first hypothesis is that the higher drying temperature (145 °C) increases the drying rate, causing the droplet surface to solidify earlier and to promote the formation of an outer shell while moisture is still present in the particle interior. Earlier surface solidification can shorten the period during which colliding droplets remain sufficiently wet and prone to coalescence, thus limiting the formation of larger particles [74]. This mechanism is therefore consistent with the observed shift toward finer particles. However, droplet coalescence was not measured directly in the present study. The second hypothesis is that accelerated drying promotes the formation of hollow particles that subsequently rupture. The internal vapor bubble trapped within the solidifying shell can cause droplets to inflate, deform, and eventually rupture, producing hollow particles, wrinkled surfaces, and other shape irregularities [75]. An increase in drying temperature also influences the evolution of the internal vapor bubble and the surrounding shell. If these hollow shells collapse, crack, or fragment during drying or subsequent handling, they could generate smaller particle fragments, thus increasing the proportion of finer particles.

The higher accessible pore volume measured by mercury intrusion porosimetry (Fig. 2) indicates that the powder produced at the higher drying temperature has a more porous structure. This observation is consistent with the second hypothesis that involves the formation of hollow or porous particles, but we cannot completely reject the first hypothesis of reduced droplet coalescence. Moreover, mercury intrusion porosimetry is a bulk characterization technique and therefore cannot determine whether the increased pore volume originates from hollow particle interiors, intragranular pores, interparticle voids, or a combination of these features. Consequently, the present measurements do not allow the dominant mechanism responsible for the finer PSD to be identified. Additional characterization, particularly high-resolution shape analysis, would help distinguish between the two hypotheses. Indeed, the SEM characterization of the spray-dried alumina particles reported in [73] reveals the coexistence of spherical, cracked, and donut-shaped particles, demonstrating that such morphological features can be observed directly and may provide valuable evidence to identify the dominant particle formation mechanism. Predominantly intact and spherical particles would support the reduced-coalescence hypothesis, whereas collapsed particles, broken shell-like structures, concave morphologies, or irregular fragments would provide evidence that shell rupture or fragmentation also contributed to the observed PSD shift.

To examine the sensitivity of DSP to the chosen simultaneous level, we perform additional analyses using simultaneously calibrated thresholds corresponding to the 90% and 99% levels (Fig. 1d). The proportion of significant bins increases from 8% (99% level) to 13.3% (95% level) and 14.6% (90% level). In a production control context, this has practical implications. Using a more conservative threshold (e.g., 99%) isolates only the most robust differences, minimizing false-positives, while a lower threshold (e.g., 90%) enhances sensitivity and may detect subtle process drifts earlier. The choice thus depends on whether the objective is tight process reproducibility or

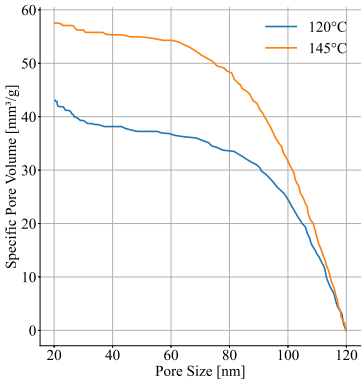


Figure 2: Specific pore volume vs. pore size of the spray-dried alumina samples at two different drying temperatures, showing increased porosity at a higher temperature.

early difference detection in quality control.

In summary, DSP quantifies both the extent and the localization of differences between the PSDs obtained under different drying temperatures. It reveals that the higher drying temperature is associated with finer and more porous particles, a pattern consistent with faster surface solidification and potentially reduced droplet coalescence or rupture of hollow particles. This coupling between statistical significance and physical interpretation underscores the utility of DSP as a diagnostic tool for linking process parameters to material attributes in spray drying.

3.2. Comparison of particles produced under identical spray-drying process conditions

A critical practical requirement for any bin-resolved comparison method is robustness against false-positive detections caused by unavoidable sampling variability and measurement noise. In industrial spray drying, even under tightly controlled conditions, repeated runs will never produce identical histograms bin-by-bin because each PSD is estimated from a finite number of detected particles. Therefore, a repeat experiment under identical operating conditions provides a natural negative-control case. Any method that marks multiple significant bins in such a comparison risks generating false alarms during routine monitoring.

To assess both the repeatability of the spray drying process and the false-positive robustness of DSP, we performed two experiments (T1 and $\mathrm{T1}_{\mathrm{repeat}}$) under identical conditions at 120 °C. The samples contain comparable particle counts ($n_1 = 2881$ and $n_2 = 3015$), providing a statistically reliable basis for comparison. The number-weighted PSDs shown in Fig. 3a exhibit a nearly complete overlap throughout the full size range, consistent with stable atomization and drying behavior for this material system.

The global Pearson statistic for this negative-control comparison is $\chi^2 = 61.6$ with a bootstrap p-value of 0.31 (> 0.05), which indicates that there is no statistically significant evidence of a global difference between the two distributions under the fitted multinomial model. This agrees with the engineering expectation that, when the atomization, drying temperature, and gas-phase conditions are kept constant, the droplet breakup and solidification pathways remain comparable, yielding closely similar particle populations.

The DSP calibration itself is also stable. The bootstrap estimate of the simultaneous significance threshold converges to $c_{0.95} = 3.15$ within 40000 iterations (Fig. 3b). Most importantly,

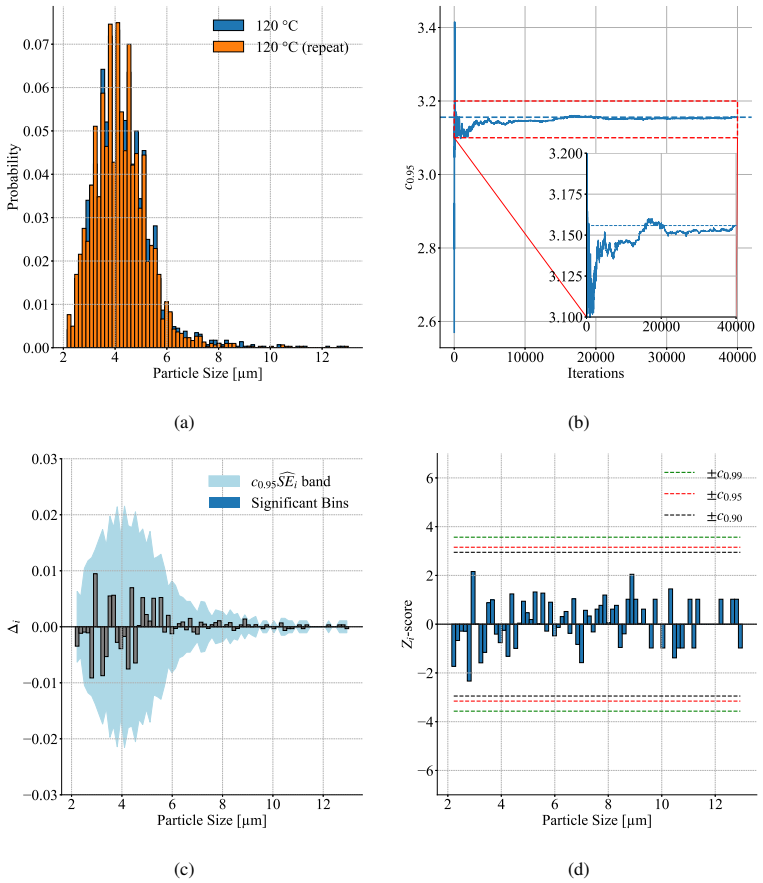
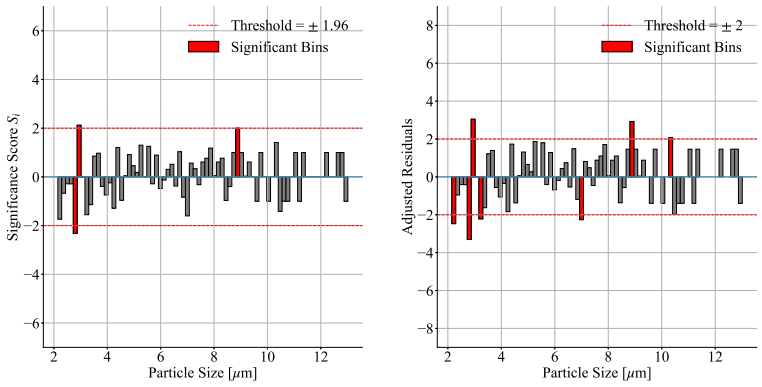


Figure 3: (a) Overlay of PSDs, (b) convergence of the simultaneous threshold $c_{0.95}$, (c) per-bin difference Δ_i with the simultaneous reference band $\pm c_{0.95}\widehat{\mathrm{SE}}_i$ (shaded blue), and (d) standardized Z_i-profile for the repeat spray drying experiments at 120 °C.

for false-positive robustness, local DSP outputs show no evidence of systematic difference. The bin-wise difference Δ_i oscillates tightly around zero, with magnitudes $|\Delta_i| < 0.01$ across the full size domain (Fig. 3c). None of the bins exceed the simultaneous $\pm c_{0.95}\widehat{\mathrm{SE}}_i$ band, and the standardized Z_i-profile remains fully contained within $\pm c_{0.95}$ (Fig. 3d). Sensitivity checks at 90% and 99% simultaneous levels similarly yield no significant bins. Collectively, these results indicate that the observed bin-to-bin differences between the repeat runs are compatible with expected sampling scatter and measurement noise, rather than suggesting a statistically significant process-induced shift.

A useful way to interpret the negative-control result is the following: since the two PSDs come from nominally identical process conditions, any local method that marks bins as significant in this comparison is, from the standpoint of process monitoring, behaving like a false alarm detector rather than identifying an intended process change. To contextualize DSP in this respect, we compare its output with two established local frameworks used for histogram-like comparisons: the normalized significance score method [53] and the adjusted residual approach [54].



(a) (b)

Figure 4: Comparison of significant bins identified by (a) Bityukov's framework and (b) Sharpe's framework for PSDs obtained from the repeat experiments at 120 °C.

Table 1: Comparison of direct bin-wise difference and three frameworks for bin-resolved PSD comparison.

Feature	**Difference of probabilities**	**Normalized significance score** [53]	**Adjusted residual** [54]	**DSP**
Variance Estimator	None	Poisson count approximation	Expected-count-based asymptotic variance	Pooled bin variance
Resampling	None	Monte Carlo	None	Parametric bootstrap
Bin Correlation	Ignored	Ignored	Ignored	Reflected in simultaneous bootstrap calibration
Multiplicity control	None	None	None	Approximate FWER control
Local significance calibration	None	Point-wise approximate	Point-wise asymptotic	Simultaneous bootstrap-calibrated

When applied to the repeat experiment sample, both alternative approaches report a non-zero set of significant bins (Fig. 4), despite the absence of statistically significant global evidence of

a difference and despite the engineering expectation of no intended process change. A likely methodological explanation is that these approaches rely on point-wise or asymptotic per-bin thresholds without simultaneously controlling the family-wise error rate across many bins and without explicitly accounting for the multinomial inter-bin dependence that naturally arises in PSD histograms. In contrast, DSP is designed specifically to reduce such false alarms by calibrating a simultaneous threshold through a parametric bootstrap that respects the multinomial model of binned PSD data. Table 1 summarizes the key methodological differences that explain this behavior.

Overall, this combined negative-control and benchmarking analysis shows that DSP preserves practical decision reliability. In the present negative-control comparison, it shows robustness against small, noise-driven bin fluctuations in repeat runs (avoiding false alarms), yet retains the ability to localize significant differences when they exceed the simultaneously calibrated uncertainty bounds. This balance between sensitivity and robustness is essential for deployment in industrial powder manufacturing workflows, where PSD measurements are inherently noisy, but process interventions are costly.

4. Conclusion

In this work, we introduce the Difference Significance Profile as a quantitative, statistically rigorous, but practically intuitive framework for comparing size distributions in process and product engineering. The framework addresses a critical gap in current practice, where visual overlays or global tests such as the χ^2 test can reveal overall differences but do not localize the specific size ranges responsible for them. By combining a multinomial-based variance formulation with parametric bootstrap calibration, DSP provides localized bin-resolved inference with bootstrap-based simultaneous error control under the fitted model. In this way, the detected differences correspond to the size classes whose observed differences are statistically significant relative to the sampling variation after accounting for multiplicity across bins.

The comparative analysis of the spray drying experiments demonstrates the complementary roles of global and local inference in powder processes. In the spray drying experiments of alumina powders, DSP identifies statistically significant redistribution of the number-weighted PSD between experiments at two different spray drying temperatures, localized primarily in the modal size range and consistent with the corresponding physical interpretation of the process. The simultaneous nature of the DSP threshold means that significant bins reflect effects that remain statistically significant after accounting for multiple comparisons, which is essential to translate PSD changes into defensible process conclusions. The negative-control repeat run further demonstrates decision reliability. DSP reported no significant bins and a non-significant global difference, indicating that the observed bin-to-bin fluctuations are compatible with sampling variability and measurement noise under the fitted model. In this negative-control comparison, the alternative methods marked non-zero significant bins under nominally identical conditions, underscoring the importance of simultaneous error control for reducing false alarms in routine monitoring and quality assurance.

In general, DSP transforms PSD comparison from a qualitative overlay into a statistically calibrated diagnostic that supports process understanding and data-informed decision-making on PSD differences. The framework is readily applicable to common particle-discrete size measurement techniques. Future work will extend the framework to mass- and volume-weighted size distribution comparisons, multivariate distribution comparisons, and integration with online monitoring systems. Such advances would further embed DSP into the workflow of modern

process analytics, enabling predictive and data-driven control of particulate systems across diverse industrial applications.

CRediT authorship contribution statement

Rahul Mitra: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization. **Lukas Fuchs**: Methodology, Writing - Review & Editing. **Ralf Ditscherlein**: Writing - Review & Editing, Visualization. **Volker Schmidt**: Writing - Review & Editing, Supervision, Project administration, Funding acquisition. **Urs Alexander Peuker**: Writing - Review & Editing, Supervision, Project administration, Funding acquisition.

Funding

The authors appreciate funding by the German Research Foundation (DFG) within SPP 2364 "Autonomous Processes in Particle Technology" under grants 504580586 and 504954383.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors thank Abhishek Prashant Singh for supporting the spray-drying experiments, Annett Kästner for providing the mercury intrusion porosimetry data, and Phillip Gräfensteiner for scientific discussion. R.M. is particularly grateful to Lisa Ditscherlein for her indispensable guidance, which was essential in shaping his understanding of the fundamentals of particle size distribution.

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

During the preparation of this work, R.M. used ChatGPT 5.5 to rephrase certain parts of the manuscript for better readability. After using this tool/service, R.M. reviewed and edited the content as needed and takes full responsibility for the content of the published article.

Data availability

The Python implementation of the DSP framework and the experimental data supporting the findings of this study are available at https://gitlab.hrz.tu-chemnitz.de/rm19lohi-at-tu-freiberg.de/dsp.git and at https://doi.org/10.25532/OPARA-1165, respectively.

Appendix A. Bin-wise covariance and bootstrap distribution

The following results illustrate how bins in a particle size distribution are statistically dependent rather than independent. Covariance heatmaps and bootstrap calibration plots make the simultaneous-inference logic of DSP visible. In particular, the covariance highlights the compositional dependence across bins, while the bootstrap distribution of the scan statistic shows why a simultaneous threshold is needed instead of an ordinary point-wise cutoff.

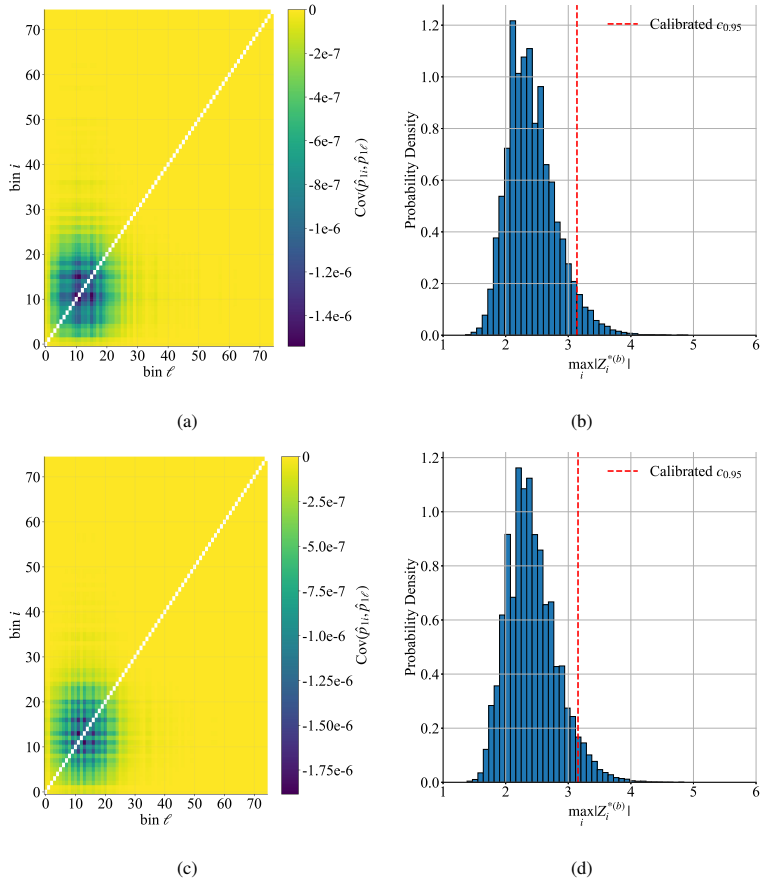


Figure A.5: Bin-wise covariance and bootstrap scan statistic distributions for the spray drying experiments. (a) Covariance $\mathrm{Cov}(\hat{p}_{1i}, \hat{p}_{1\ell})$ for the comparison between 120 °C and 145 °C; (b) bootstrap distribution of $\max_i |Z_i^{*(b)}|$ for this comparison, showing the calibrated threshold $c_{0.95} = 3.14$; (c) $\mathrm{Cov}(\hat{p}_{1i}, \hat{p}_{1\ell})$ for the repeat experiment at 120 °C; (d) corresponding bootstrap distribution of $\max_i |Z_i^{*(b)}|$, again with calibrated threshold $c_{0.95} = 3.15$.

For spray drying experiments at 120 °C and 145 °C, Fig. A.5a shows that the covariance is concentrated almost entirely in the size regime $i, \ell \lesssim 25$, with values around -1.4×10^{-6}. Beyond about the 30$^{\text{th}}$ bin, both estimated bin probabilities $\hat{p}_{1i}$ and $\hat{p}_{2i}$ are close to zero, so the corresponding covariances become negligible. This localization indicates that sampling uncertainty and cross-bin dependence are dominated by the size range where most of the population is concentrated and where the PSD is most sensitive to changes in operating conditions.

The bootstrap distribution of $\max_i |Z_i^{*(b)}|$ for this sample (Fig. A.5b) peaks near 2.5 and extends beyond 4.0, resulting in a calibrated threshold $c_{0.95} = 3.14$. This shows that, even for experimental data affected by heterogeneity and measurement noise, the distribution of the maximum standardized fluctuation is substantially wider than a point-wise reference would suggest. The resulting threshold is therefore consistent with approximate family-wise error control at the nominal 5% level under the fitted multinomial model.

For the negative-control experiments (Fig. A.5c), both the covariance and the calibrated threshold are very similar to those of the original spray drying comparison. No bins exceed the simultaneous significance threshold, and the global Pearson χ^2 test remains non-significant. Together, these results support the interpretation of reproducible measurements and stable process behavior under nominally identical processing conditions.

These analyses support the theoretical foundations of DSP. The consistently non-diagonal covariance matrices across samples show that bin-wise probability estimates are correlated through the compositional structure of the PSD. The bootstrap-based calibration of the scan statistic explicitly accounts for this dependence, providing a single global threshold for model-based simultaneous inference. Without this calibration, point-wise thresholds such as 1.96 would substantially inflate the probability of declaring at least one bin significant by chance alone when many bins are examined simultaneously, especially in regions with strong bin dependence.

References

[1] M. Donoso, P. A. Robles, E. D. Gálvez, L. A. Cisternas, Particle size effect on the efficient use of water and energy in mineral concentration processes, Industrial and Engineering Chemistry Research 52 (49) (2013) 17686–17690. doi:10.1021/ie402099n.

[2] W. Guo, K. Guo, Y. Xing, X. Gui, A comprehensive review on evolution behavior of particle size distribution during fine grinding process for optimized separation purposes, Mineral Processing and Extractive Metallurgy Review 47 (1) (2026) 1–20. doi:10.1080/08827508.2024.2410287.

[3] D. Bourell, B. Stucker, A. Spierings, N. Herres, G. Levy, Influence of the particle size distribution on surface quality and mechanical properties in AM steel parts, Rapid Prototyping Journal 17 (3) (2011) 195–202. doi:10.1108/13552541111124770.

[4] A. Kaas, C. Wilke, A. Vanderbruggen, U. A. Peuker, Influence of different discharge levels on the mechanical recycling efficiency of lithium-ion batteries, Waste Management 172 (2023) 1–10. doi:10.1016/j.wasman.2023.08.042.

[5] K. Horváth, D. Lukács, A. Sepsey, A. Felinger, Effect of particle size distribution on the separation efficiency in liquid chromatography, Journal of Chromatography A 1361 (2014) 203–208. doi:10.1016/j.chroma.2014.08.017.

[6] D. Barone, E. Loth, P. Snyder, Influence of particle size on inertial particle separator efficiency, Powder Technology 318 (2017) 177–185. doi:10.1016/j.powtec.2017.04.044.

[7] Y. Zhao, J. W. Chew, Effect of lognormal particle size distributions on particle spreading in additive manufacturing, Advanced Powder Technology 32 (4) (2021) 1127–1144. doi:10.1016/j.apt.2021.02.019.

[8] O. Voigt, T. Buchwald, L. Ditscherlein, R. Ditscherlein, U. A. Peuker, Flowability, shape and structure characterization of recycled particles originating from electro discharge machining of H11 alloy, Advanced Powder Technology 36 (11) (2025) 105064. doi:10.1016/j.apt.2025.105064.

[9] A. Stunda-Zujeva, Z. Irbe, L. Berzina-Cimdina, Controlling the morphology of ceramic and composite powders obtained via spray drying – A review, Ceramics International 43 (15) (2017) 11543–11551. doi:10.1016/j.ceramint.2017.05.023.

[10] C.-L. Lin, T.-H. Peng, W.-J. Wang, Effect of particle size distribution on agglomeration/defluidization during fluidized bed combustion, Powder Technology 207 (1) (2011) 290–295. doi:10.1016/j.powtec.2010.11.010.

[11] S. C. Das, S. R. B. Behara, D. A. Morton, I. Larson, P. J. Stewart, Importance of particle size and shape on the tensile strength distribution and de-agglomeration of cohesive powders, Powder Technology 249 (2013) 297–303. doi:10.1016/j.powtec.2013.08.034.

[12] J. Nicklas, U. Peuker, Agglomeration of fine hydrophobic particles: 1D and 2D characterization by dynamic image analysis of in-line probe data, Powder Technology 426 (2023) 118685. doi:10.1016/j.powtec.2023.118685.

[13] N. Eiermann, O. Furat, J. Nicklas, U. A. Peuker, V. Schmidt, Quantitative characterization of hydrophobic agglomeration at different mixing intensities using a copula-based probabilistic modeling approach, Powder Technology 460 (2025) 120907. doi:10.1016/j.powtec.2025.120907.

[14] R. I. Mackie, R. Bai, The role of particle size distribution in the performance and modelling of filtration, Water Science and Technology 27 (10) (1993) 19–34. doi:10.2166/wst.1993.0199.

[15] J. S. Chang, S. Vigneswaran, J. K. Kandasamy, L. J. Tsai, Effect of pore size and particle size distribution on granular bed filtration and microfiltration, Separation Science and Technology 43 (7) (2008) 1771–1784. doi:10.1080/01496390801974605.

[16] N. Barth, C. Schilde, A. Kwade, Influence of particle size distribution on micromechanical properties of thin nanoparticulate coatings, Physics Procedia 40 (2013) 9–18. doi:10.1016/j.phpro.2012.12.002.

[17] Y. Wu, L. F. Francis, Effect of particle size distribution on stress development and microstructure of particulate coatings, Journal of Coatings Technology and Research 14 (2) (2017) 455–465. doi:10.1007/s11998-016-9866-5.

[18] J. M. Friebel, R. Ditscherlein, L. Ditscherlein, U. A. Peuker, Three-dimensional characterization of dry particle coating structures originating from the mechano-fusion process, Microscopy and Microanalysis 30 (2) (2024) 179–191. doi:10.1093/mam/ozae009.

[19] H. Ding, Z. Ouyang, X. Zhang, S. Zhu, The effects of particle size on flameless combustion characteristics and nox emissions of semi-coke with coal preheating technology, Fuel 297 (2021) 120758. doi:10.1016/j.fuel.2021.120758.

[20] H. Yan, B. Nie, F. Kong, Y. Liu, P. Liu, Y. Wang, Z. Chen, F. Yin, J. Gong, S. Lin, X. Wang, Y. Hou, Experimental investigation of coal particle size on the kinetic properties of coal oxidation and spontaneous combustion limit parameters, Energy 270 (2023) 126890. doi:10.1016/j.energy.2023.126890.

[21] A. van Schaik, M. Reuter, K. Heiskanen, The influence of particle size reduction and liberation on the recycling rate of end-of-life vehicles, Minerals Engineering 17 (2) (2004) 331–347. doi:10.1016/j.mineng.2003.09.019.

[22] M. Li, J. Zhang, W. Song, D. M. Germain, Recycling of crushed waste rock as backfilling material in coal mine: effects of particle size on compaction behaviours, Environmental Science and Pollution Research 26 (9) (2019) 8789–8797. doi:10.1007/s11356-019-04379-9.

[23] S. Zong, C. Chang, P. Rem, A. T. Gebremariam, F. Di Maio, Y. Lu, Research on the influence of particle size distribution of high-quality recycled coarse aggregates on the mechanical properties of recycled concrete, Construction and Building Materials 465 (2025) 140253. doi:10.1016/j.conbuildmat.2025.140253.

[24] C. Servais, R. Jones, I. Roberts, The influence of particle size distribution on the processing of food, Journal of Food Engineering 51 (3) (2002) 201–208. doi:10.1016/S0260-8774(01)00056-5.

[25] T. Chen, M. Zhang, B. Bhandari, Z. Yang, Micronization and nanosizing of particles for an enhanced quality of food: A review, Critical Reviews in Food Science and Nutrition 58 (6) (2018) 993–1001. doi:10.1080/10408398.2016.1236238.

[26] Z. Štirbanović, J. Sokolović, I. Marković, S. Đorđievski, The effect of degree of liberation on copper recovery from copper-pyrite ore by flotation, Separation Science and Technology 55 (17) (2020) 3260–3273. doi:10.1080/01496395.2019.1676260.

[27] S. Sahu, N. Devi, Two-step leaching of spent lithium-ion batteries and effective regeneration of critical metals and graphitic carbon employing hexuronic acid, RSC Advances 13 (2023) 7193–7205. doi:10.1039/D2RA07926G.

[28] M. Xue, J. Li, Z. Xu, Environmental friendly crush-magnetic separation technology for recycling metal-plated plastics from end-of-life vehicles, Environmental Science & Technology 46 (5) (2012) 2661–2667. doi:10.1021/es202886a.

[29] C. Bin, Y. Yi, S. Zhicheng, W. Qiang, A. Abdelkader, A. R. Kamali, D. Montalvão, Effects of particle size on the separation efficiency in a rotary-drum eddy current separator, Powder Technology 410 (2022) 117870. doi:10.1016/j.powtec.2022.117870.

[30] N. Vlachos, I. T. Chang, Investigation of flow properties of metal powders from narrow particle size distribution to polydisperse mixtures through an improved Hall-flowmeter, Powder Technology 205 (1) (2011) 71–80. doi:10.1016/j.powtec.2010.08.067.

[31] Z. Young, M. Qu, M. M. Coday, Q. Guo, S. M. H. Hojjatzadeh, L. I. Escano, K. Fezzaa, L. Chen, Effects of particle size distribution with efficient packing on powder flowability and selective laser melting process, Materials 15 (3) (2022). doi:10.3390/ma15030705.

[32] S. Chandra, C. Wang, S. B. Tor, U. Ramamurty, X. Tan, Powder-size driven facile microstructure control in powder-fusion metal additive manufacturing processes, Nature Communications 15 (1) (2024). doi:10.1038/s41467-024-47257-w.

[33] H. Rumpf, Über die Eigenschaften von Nutzstäuben, Staub-Reinhaltung der Luft 27 (1) (1967) 3–13.

[34] W. Peukert, Material properties in fine grinding, International Journal of Mineral Processing 74 (2004) S3–S17. doi:10.1016/j.minpro.2004.08.006.

[35] R. Ditscherlein, O. Furat, E. Löwer, R. Mehnert, R. Trunk, T. Leißner, M. J. Krause, V. Schmidt, U. A. Peuker, PARROT: A pilot study on the open access provision of particle-discrete tomographic datasets, Microscopy and Microanalysis 28 (2) (2022) 350–360. doi:10.1017/S143192762101391X.

[36] L. Zhao, S. Zhang, D. Huang, X. Wang, A digitalized 2D particle database for statistical shape analysis and discrete modeling of rock aggregate, Construction and Building Materials 247 (2020) 117906. doi:10.1016/j.conbuildmat.2019.117906.

[37] H. G. Merkus, Particle size, size distributions and shape, Springer, 2009, pp. 13–42. doi:10.1007/978-1-4020-9016-5_2.

[38] D. R. Handwerk, P. D. Shipman, C. B. Whitehead, S. Özkar, R. G. Finke, Particle size distributions via mechanism-enabled population balance modeling, The Journal of Physical Chemistry C 124 (8) (2020) 4852–4880. doi:10.1021/acs.jpcc.9b11239.

[39] I. Gonz'alez Tejada, P. Antolin, Use of machine learning for unraveling hidden correlations between particle size distributions and the mechanical behavior of granular materials, Acta Geotechnica 17 (5) (2022) 1443–1461. doi:10.1007/s11440-021-01420-5.

[40] H. Hatzantonis, A. Goulas, C. Kiparissides, A comprehensive model for the prediction of particle-size distribution in catalyzed olefin polymerization fluidized-bed reactors, Chemical Engineering Science 53 (18) (1998) 3251–3267. doi:10.1016/S0009-2509(98)00122-5.

[41] H. Abdulhussain, M. Thompson, Predicting the particle size distribution in twin screw granulation through acoustic emissions, Powder Technology 394 (2021) 757–766. doi:10.1016/j.powtec.2021.08.089.

[42] R. L. Scheaffer, A. Watkins, M. Gnanadesikan, J. A. Witmer, Estimating the difference between two proportions, Springer, 1996, pp. 134–138. doi:10.1007/978-1-4757-3843-8_30.

[43] K. Pearson, On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science 50 (302) (1900) 157–175. doi:10.1080/14786440009463897.

[44] R. C. Dahiya, On the Pearson chi-squared goodness-of-fit test statistic, Biometrika 58 (3) (1971) 685–686. doi:10.2307/2334410.

[45] S. D. Bolboacă, L. Jäntschi, A. F. Sestraş, R. E. Sestraş, D. C. Pamfil, Pearson-Fisher chi-square statistic revisited, Information 2 (3) (2011) 528–545. doi:10.3390/info2030528.

[46] F. J. Massey Jr, The Kolmogorov-Smirnov test for goodness of fit, Journal of the American statistical Association 46 (253) (1951) 68–78. doi:10.1080/01621459.1951.10500769.

[47] K. Matusita, Decision rules, based on the distance, for problems of fit, two samples, and estimation, The Annals of Mathematical Statistics 26 (4) (1955) 631–640. doi:10.1214/aoms/1177728422.

[48] S. Verdu, Total variation distance and the distribution of relative information, in: Information Theory and Applications Workshop (ITA), IEEE, 2014, p. 1–3. doi:10.1109/ita.2014.6804281.

[49] A. Gretton, K. M. Borgwardt, M. J. Rasch, B. Schölkopf, A. Smola, A kernel two-sample test, Journal of Machine Learning Research 13 (25) (2012) 723–773.
URL http://jmlr.org/papers/v13/gretton12a.html

[50] F. Nielsen, On the Jensen–Shannon symmetrization of distances relying on abstract means, Entropy 21 (5) (2019). doi:10.3390/e21050485.

[51] V. M. Panaretos, Y. Zemel, Statistical aspects of Wasserstein distances, Annual Review of Statistics and Its Application 6 (1) (2019) 405–431. doi:10.1146/annurev-statistics-030718-104938.

[52] M. B. Wilk, R. Gnanadesikan, Probability plotting methods for the analysis of data, Biometrika 55 (1) (1968) 1–17. doi:10.1093/biomet/55.1.1.

[53] S. Bityukov, N. Krasnikov, A. Nikitenko, V. Smirnova, A method for statistical comparison of histograms (2013). arXiv:1302.2651, doi:10.48550/arXiv.1302.2651.

[54] D. Sharpe, Chi-square test is statistically significant: Now what?, Practical Assessment, Research, and Evaluation 20 (1) (2015) 8. doi:10.7275/tbfa-x148.

[55] J. J. Goeman, A. Solari, Multiple testing for exploratory research, Statistical Science 26 (4) (2011). doi:10.1214/11-sts356.

[56] T. Dickhaus, J. Stange, Multiple point hypothesis test problems and effective numbers of tests for control of the family-wise error rate, Calcutta Statistical Association Bulletin 65 (1-4) (2013) 123–144. doi:10.1177/0008068320130108.

[57] B. Efron, R. J. Tibshirani, An introduction to the bootstrap, Chapman and Hall, 1994. doi:10.1201/9780429246593.

[58] T. Matsuyama, An application of bootstrap method for analysis of particle size distribution, Advanced Powder Technology 29 (6) (2018) 1404–1408. doi:10.1016/j.apt.2018.03.002.

[59] E. Schach, M. Buchmann, R. Tolosana-Delgado, T. Leißner, M. Kern, K. Gerald van den Boogaart, M. Rudolph, U. A. Peuker, Multidimensional characterization of separation processes – Part 1: Introducing kernel methods and entropy in the context of mineral processing using sem-based image analysis, Minerals Engineering 137 (2019) 78–86. doi:10.1016/j.mineng.2019.03.026.

[60] T. Matsuyama, Parametric and non-parametric evaluation of conversion of number-based particle size distribution to mass-based distribution, Advanced Powder Technology 35 (9) (2024) 104594. doi:10.1016/j.apt.2024.104594.

[61] C. Evans, T. Napier-Munn, Estimating error in measurements of mineral grain size distribution, Minerals Engineering 52 (2013) 198–203, process Mineralogy. doi:10.1016/j.mineng.2013.09.005.

[62] R. Mariano, C. Evans, Error analysis in ore particle composition distribution measurements, Minerals Engineering 82 (2015) 36–44. doi:10.1016/j.mineng.2015.06.001.

[63] D. Freedman, P. Diaconis, On the histogram as a density estimator: L2 theory, Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete 57 (4) (1981) 453–476. doi:10.1007/BF01025868.

[64] D. W. Scott, On optimal and data-based histograms, Biometrika 66 (3) (1979) 605–610. doi:10.1093/biomet/66.3.605.

[65] H. A. Sturges, The choice of a class interval, Journal of the American Statistical Association 21 (153) (1926) 65–66. doi:10.1080/01621459.1926.10502161.

[66] T. Jiang, S. Luo, D. Wang, Y. Li, Y. Wu, L. He, G. Zhang, A new bin size index method for statistical analysis of multimodal datasets from materials characterization, Scientific Reports 13 (1) (2023) 10915. doi:10.1038/s41598-023-37969-2.

[67] A. Agresti, Categorical data analysis, Wiley & Sons, Inc., 2002. doi:10.1002/0471249688.

[68] C. Elkan, Clustering documents with an exponential-family approximation of the Dirichlet compound multinomial distribution, in: Proceedings of the 23rd International Conference on Machine learning - ICML '06, ACM Press, 2006, p. 289–296. doi:10.1145/1143844.1143881.

[69] N. Zamzami, N. Bouguila, Sparse count data clustering using an exponential approximation to generalized dirichlet multinomial distributions, IEEE Transactions on Neural Networks and Learning Systems 33 (1) (2022) 89–102. doi:10.1109/tnnls.2020.3027539.

[70] C. A. W. Glas, Maximum-likelihood estimation, Chapman and Hall, 2017, p. 197–217. doi:10.1201/b19166-11.

[71] Z. Holcomb, Fundamentals of descriptive statistics, Routledge, 2016. doi:10.4324/9781315266510.

[72] P. Good, Permutation, parametric, and bootstrap tests of hypotheses, Springer, 2005. doi:10.1007/b138696.

[73] R. Mitra, L. Fuchs, O. Furat, Y. Sinnwell, S. Antonyuk, V. Schmidt, U. A. Peuker, The impact of the feed rate and the binder concentration on the morphology of spray-dried alumina–polymer nanocomposites, Processes 13 (6) (2025). doi:10.3390/pr13061643.

[74] P. Biswas, D. Sen, S. Mazumder, C. B. Basak, P. Doshi, Temperature mediated morphological transition during drying of spray colloidal droplets, Langmuir 32 (10) (2016) 2464–2473. doi:10.1021/acs.langmuir.5b04171.

[75] J. P. Hecht, C. J. King, Spray drying: Influence of developing drop morphology on drying rates and retention of volatile substances. 1. Single-drop experiments, Industrial & Engineering Chemistry Research 39 (6) (2000) 1756–1765. doi:10.1021/ie9904652.