Analysis of Multivariate Industrial Process Data for Quality Pattern Detection in Masterbatch Coloring

Leonard Hisgen[1,2], Thomas Kubik[1], Sascha Garbe[3], Jens Fischer[4],
Alexander Kloes[1], Benjamin Iniguez[2], Mike Schwarz[1]
[1] NanoP, THM University of Applied Sciences, 35390 Giessen, Germany
[2] DEEEA, Universitat Rovira i Virgili, 43002 Tarragona, Spain
[3] G.E. HABICH'S SÖHNE GmbH & Co. KG, 34359 Reinhardshagen, Germany
[4] superus Datenmanagement GmbH, 35099 Burgwald, Germany
Email: leonard.hisgen@nanop.thm.de

***Abstract*—This paper analyzes clustering techniques for identifying quality-related structures in multivariate industrial process data, demonstrated through a case study on masterbatch data from the coloring industry. The dataset consists of batches labeled as successful or unsuccessful based on quality inspection results. The proposed approach combines dimensionality reduction, clustering, and statistical analysis to examine relationships between ingredient composition and quality outcomes. Multiple clustering methods are compared to identify characteristic differences in ingredient usage. The results demonstrate that clustering methods reveal distinct patterns in ingredient composition associated with quality outcomes.**

***Keywords*—clustering, industrial coloring, masterbatch, multivariate process data, unsupervised learning**

Summary

Modern industrial production processes generate large volumes of multivariate data that contain valuable information about process behavior and product quality. However, extracting meaningful insights from such high-dimensional data remains challenging using conventional analysis methods [1]. In industrial coloring processes, variations in ingredient composition and production conditions can significantly influence batch quality, and failed batches require rework, resulting in additional time and cost [2]. While clustering methods are widely used in industrial data analysis, their application to multicomponent recipe data in coloring and masterbatch production remains comparatively limited [1].

This work investigates clustering methods for identifying quality-related structures in multivariate industrial masterbatch data. The analysis is based on a confidential production dataset containing ingredient compositions and quality inspection results. After preprocessing, dimensionality reduction methods are used to analyze global and local data structures, while clustering approaches including k-means, DBSCAN, and hierarchical clustering are compared using internal and external evaluation metrics. Statistical analyses are then applied to interpret cluster-specific ingredient differences.

The results show that global structures in the full dataset are mainly driven by differences between product variants, whereas practically relevant and interpretable quality-related patterns emerge at the level of individual product variants. Across the analyzed variants, hierarchical clustering and k-means provide the most stable results, while DBSCAN captures meaningful structures in selected cases but is more sensitive to data characteristics and parameter settings. A representative case study further illustrates that local structures within a product variant can be associated with differences in quality outcomes (see Fig. 1).

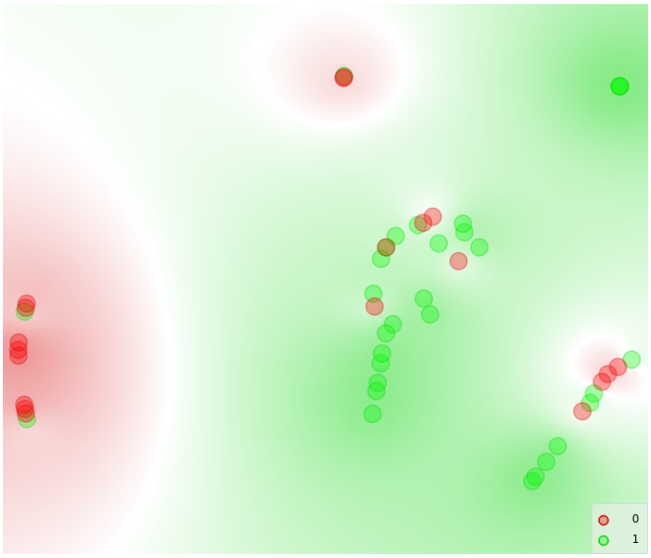

Fig. 1. t-SNE visualization of a grey color product variant dataset. Red markers indicate failed quality tests and green markers indicate successful quality tests.

Overall, the study demonstrates that clustering methods are suitable for exploratory analysis of multivariate industrial process data and can support the identification of composition-quality relationships. The presented work serves as a prestudy for data-driven and explainable AI approaches in industrial coloring processes.

References

[1] C. Ji and W. Sun, "A review on data-driven process monitoring methods: Characterization and mining of industrial data," *Processes*, vol. 10, no. 2, p. 335, Feb. 2022, doi: 10.3390/pr10020335.
[2] G. Etzrodt, *Industrial Coloration of Plastics: Pigments, Dyestuffs, Fillers, and Nanomaterials*. München: Hanser, 2022.